A NONINVASIVE, IMAGE-BASED SMARTPHONE APP FOR DIAGNOSING ANEMIA

A Dissertation
Presented to
The Academic Faculty

by

Robert G. Mannino

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Wallace H. Coulter Department of Biomedical Engineering

Georgia Institute of Technology and Emory University
May 2018

A NONINVASIVE, IMAGE-BASED SMARTPHONE APP FOR DIAGNOSING ANEMIA

Approved by:

Dr. Wilbur A. Lam, Advisor
College of Engineering
Wallace H. Coulter Department of
Biomedical Engineering
Georgia Institute of Technology
and Emory University

Dr. Peng Qiu
College of Engineering
Wallace H. Coulter Department of
Biomedical Engineering
Georgia Institute of Technology
and Emory University

Dr. Gari D. Clifford
College of Engineering
Wallace H. Coulter Department of
Biomedical Engineering
Georgia Institute of Technology
and Emory University

Dr. Jeanne Boudreaux
Aflac Cancer Center
& Blood Disorder Center
Children's Healthcare of Atlanta

Dr. Lee Cooper
College of Engineering
Wallace H. Coulter Department of
Biomedical Engineering
Georgia Institute of Technology
and Emory University

Date Approved: [March 05, 2018]

ACKNOWLEDGEMENTS

I would like to start by thanking my family for supporting me throughout my academic career. I would like to thank my wife, Nicole Mannino, for her unwavering support and encouragement every step of the way. I would like to thank my parents, Joy and Steve Mannino, for believing in me no matter what and for always encouraging me to push myself and never quit. I would like to thank my brothers, Kevin and Patrick Mannino, for their support and for motivating me to be a better person.

I would like to thank my advisor, Dr. Wilbur Lam, for taking me into his lab as an undergraduate researcher, sparking my interest in research, and providing immeasurable support, guidance, and mentorship throughout my research career. I would not have chosen to pursue graduate school with any other advisor and, thus, would certainly not be in this position today. I would like to thank my colleagues in the Lam Lab – Dr. Elaissa Trybus Hardy, Dr. Yongzhi Qiu, Dr. Sarah Mitchell, Dr. Marcus Carden, Dr. Margo Rollins, Dr. Hope Gole, Dr. Byungwook Ahn, Dr. Jordan Ciciliano, Dr. Caroline Hansen, Erika Tyburski, Meredith Fay, Jessica Lin, John Nicosia, Kendall Williams, and Mo Ohinowo – for their support and helpful conversations.

I would like to thank my thesis committee members for their ongoing support of my project. Thank you to Dr. Gari Clifford for providing detailed technical advice that helped this project get off the ground. Thank you to Dr. Lee Cooper and Dr. Peng Qiu for their valuable insights and critiques. I would like to especially thank Dr. Jeanne Boudreaux, my pediatric hematologist/thesis committee member for caring for me as a child, sparking my interest in science, and introducing me to my advisor.

Finally, I would like to thank my friends, who without their support and friendship, I could never have made it this far. I would like to thank my best friend, Davin Quarshie, for his continued friendship, expert speechwriting, and willingness to debate any issue. I would like to thank Dr. Reginald Tran, for being an exemplary role model scientist, powerlifter, and friend. I would like to thank Dr. David Myers, for being my first research mentor, and guiding me as a researcher and as a person. I would like to thank Yumiko Sakurai for her brilliant advice regarding experimental biology, as well as sharing with me her love of nature, photography, and knitting. I would also like to thank the boys – Clark Kerr, Muaz Rushdi, Preston Mayo, Adam Swoboda, Grayson Baum and Victor Baldwin –for making sure that I was always having fun, regardless of what grad school was throwing at me.

Thank you all. This work would not have been possible without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

API	Application Programming Interface
CBC	Complete Blood Count
CHOA	Children's Healthcare of Atlanta
CI	Confidence Interval
EMR	Electronic Medical Records
FDA	Food and Drug Administration
GUI	Graphical User Interface
Hgb	Hemoglobin
HIPAA	Health Insurance Portability and Accountability Act
IDE	Integrated Development Environment
IRB	Institutional Review Board
mHealth	Mobile Health
POC	Point of Care
RBC	Red Blood Cell
RGB	Red, Green, Blue color space
ROI	Region of Interest

SUMMARY

Smartphone-based telehealth is steadily transforming the delivery of medical care worldwide, moving diagnosis of disease from the clinic to the home to potentially anywhere in the globe. Smartphone images alone have recently been used by physicians to remotely diagnose a myriad of diseases. However, smartphone telehealth approaches have yet to non-invasively replace blood-based testing, which remains a major cornerstone of disease diagnosis in modern medicine. While the addition of specialized smartphone attachments and supplemental calibration tools may enable point-of-care diagnosis and analysis of tissue and bodily fluid samples, the additional burden of blood and/or tissue sample collections combined with the additional cost and inconvenience associated with this equipment, prevents worldwide use of these potentially disruptive approaches. Therefore, a smartphone-based system, requiring nothing other than the smartphones native technology and capable of non-invasively replacing blood-based diagnostics, would transform the very nature of telehealth and the delivery of healthcare worldwide. Towards that end, I specifically focused on anemia, a potentially life-threatening disorder characterized by low blood hemoglobin (Hgb) levels that affects approximately 2 billion people worldwide. Despite the high prevalence of anemia, all existing diagnostic approaches to measure Hgb require specialized equipment and represent tradeoffs between invasiveness, accuracy, infrastructure needs, and expense. Aside from being cost-prohibitive, the necessary invasive blood sampling to measure Hgb levels causes discomfort and trauma in younger pediatric patients. By examining clinical pallor, a common symptom of anemia, I developed a methodology that

quantitatively analyzes patient-sourced photos using smartphone-based algorithms to enable a noninvasive, accurate, and accessible anemia diagnostic. Here, a patient simply takes a picture of their fingernail beds using their smartphone, and the image analysis algorithm analyzes color data and image metadata to measure the corresponding Hgb level. By quantifying clinical pallor, this system non-invasively measures Hgb levels to within a clinically significant and well accepted margin of error (±2.6 g/dL) of the gold standard Hgb measurement tool with a sensitivity and specificity of 0.90 and 0.82, respectively, of predicting anemia (defined as Hgb < 11.0g/dL) in 100 pediatric patients at Children's Healthcare of Atlanta with anemia of any etiology mixed with healthy subjects. Furthermore, this algorithm can be personally calibrated to achieve and accuracy of ±0.9 g/dL, a degree of accuracy which, upon further testing, may enable this technique to replace traditional blood-based testing. This algorithm has been implemented into a smartphone app that is capable of outperforming trained hematologists in physical examination-based Hgb measurement. Overall, this technology has the capability to change the treatment paradigm for anemia as patients no longer need to visit a clinic to monitor their hemoglobin. In this thesis, I discuss the development of this image analysis algorithm and the implementation of the algorithm into a smartphone app.

CHAPTER 1. INTRODUCTION

1.1 Motivation

Anemia, characterized by low blood hemoglobin (Hgb) levels[1], is a global public health problem that affects an estimated 1 in 3 people worldwide. Anemia is caused by a variety of etiologies, including chronic hematological conditions that disproportionately affect underserved populations, such as sickle cell disease and β-thalassemia[2,3]. Patients suffering from chronic anemia require frequent monitoring of their Hgb levels to track the progression of their disease and inform their treatment strategy. The gold standard Hgb measurement test is the complete blood count (CBC) which typically requires a blood draw and trained laboratory technicians as well as expensive analytical equipment and reagents[4]. Despite the high prevalence of anemia, there is currently no noninvasive, inexpensive, and accurate Hgb assessment technology available that enables patients with chronic conditions to better self-manage their disease[5-7].

To that end, advancements in point-of-care (POC) technology are dramatically and cost-effectively improving clinical outcomes and quality of life of patients with chronic diseases by facilitating patient-provider communication and enabling self-management[8]. Smartphones offer an ideal POC platform, as they are already distributed worldwide and in the hands of billions of users[9]. Recently, patients have used smartphones to capture and transmit images to physicians for remote diagnosis, such as in cases of ear infections or melanoma[10,11]. Moreover, with customized attachments, smartphones have been widely adapted to perform POC diagnostic analysis on body tissue and fluid samples[12]. However, there are currently no smartphone-based approaches

that can non-invasively replace blood-based diagnostics (e.g. Hgb measurement), a major cornerstone in the practice of modern medicine. A patient-operated, home-based method to measure Hgb as a self-test would dramatically improve the quality of life of chronic anemia patients, as this would obviate the need for frequent, time-consuming, and expensive clinic visits to measure Hgb levels.

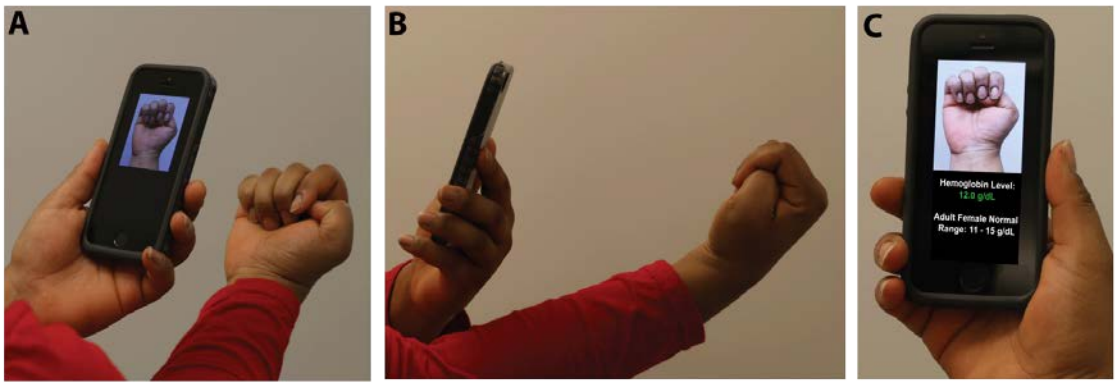

Figure 1: This smartphone app enables non-invasive, quantitative self-measurement of patient blood Hgb levels. A) A patient takes a photo of their fingernail beds. **B)** Using the native hardware on the smartphone, without needing any attachments or calibration equipment, **(C)** the image analysis algorithm quantitatively measures Hgb levels.

To address these challenges, I have developed and clinically validated a novel, non-invasive smartphone-based technology capable of accurately measuring blood Hgb levels using only patient-sourced images and the phone's native hardware without additional attachments or calibration tools. Using this smartphone app, a patient simply takes an image of their fingernail beds, where a custom, novel, image analysis algorithm calculates that patient's Hgb level based on the pallor of their fingernail beds (Figure 1). Clinicians have used physical examination of the fingernails, conjunctiva, and palmar creases for qualitative assessment of anemia, and several groups have semi-quantitatively characterized anemia using these clinical findings[13-17]. Combining this clinically

established utility of physical examination in anemia diagnosis with the imaging capabilities of smartphones, this image analysis algorithm relates clinical pallor of the fingernail beds from patient images to quantitative Hgb levels.

1.2 Research Objectives and Specific Aims

The overall goals of this project were to: (Aim1) determine the feasibility of developing an image analysis algorithm that can accurately predict Hgb values from smartphone images of subject's fingernails, (Aim 2) develop the anemia prediction algorithm and validate with a large scale clinical assessment (n = 337), and (Aim 3) perform usability testing and incorporate this algorithm into a translational, easy to use, and inexpensive smartphone application. The incorporation of this image analysis algorithm into a mobile platform in the form of an app ultimately facilitates patient self-management by allowing chronic anemia patients to measure their own Hgb levels, enabling more effective patient-provider communication, as well as improving access to healthcare especially in low resource settings.

1.2.1 Aim 1: Determine the feasibility of developing an image analysis algorithm that can accurately predict Hgb values from smartphone images of subjects fingernails

In this aim, I showed that it is possible to quantitatively correlate the clinical pallor in a patients fingernail beds to the patient's blood Hgb concentration. Images of fingernail beds were taken using a smartphone and multiple parameters were extracted from the images. These parameters were then correlated with actual Hgb levels to determine their relationships to blood

Hgb level. Accurate Hgb levels used to evaluate the regression analysis were obtained via a blood draw and subsequent measurement via a CBC, the gold standard for measuring anemia.

1.2.2 Aim 2: Validate and improve the anemia monitoring algorithm with a large scale clinical assessment

In this aim, I conducted a large scale validation study at Children's Healthcare of Atlanta (CHOA), to evaluate the accuracy of the Hgb measurement algorithm. Multi-linear regression was used to correlate image parameters of subject's fingernail beds with gold-standard measured Hgb. Once an algorithm that accurately predicts Hgb values was developed, the algorithm was validated in a large scale clinical assessment. This involved taking images of hospitalized patient's fingernail beds, and comparing the Hgb results of the prediction algorithm with results from the gold standard obtained during the course of the patient's standard treatment protocol. Additionally, chronically anemic patients whose Hgb levels fluctuate were monitored over time to assess the utility of this prediction algorithm as a self-monitoring tool that can be tailored to an individual's physical characteristics. Data collected from this clinical study was then used to iteratively improve the prediction algorithm

1.2.3 Perform usability testing and incorporate this algorithm into a translational, easy to use, and inexpensive smartphone application

In this aim, I incorporated the Hgb measurement algorithm in smartphone apps running on both the iOS and Android operating systems. Testing was conducted to determine the testing conditions under which the app will work. This was done by comparing the accuracy of the app when testing was conducted using multiple different smartphones as well as varying background lighting condition of the imaging location.

1.3 Significance

Anemia, characterized by low blood Hgb levels, is the world's most common blood disorder, afflicting nearly two billion people worldwide[1]. Due to numerous causes including malnutrition (e.g. iron deficiency), genetic diseases (e.g. sickle cell disease and β-thalassemia), infections/inflammatory disorders, cancer chemotherapy, and renal failure, anemia can be chronic and potentially life-threatening[2,18-22]. Therefore, at-risk individuals, especially those with chronic anemia, require frequent Hgb level measurements to monitor their disease and dictate their treatment plan. For example, sickle cell disease, β-thalassemia, and marrow failure often require chronic transfusions, the amount and timing of which are determined by Hgb levels. Also, patients with autoimmune anemia require frequent Hgb measurements to determine their medication type and dosage. Currently, the gold-standard Hgb diagnostic is the CBC[4], which requires a costly patient visit to a clinic/hospital or commercial lab, expensive analytical equipment, as well as trained phlebotomists/technicians. Often the transportation logistics alone to obtain these labs are cost-prohibitive to patients, especially those who are members of underserved populations, and the necessary invasive blood sampling to measure Hgb levels causes discomfort and trauma in younger pediatric patients, a known demographic at risk for anemia[23]. Due to the inconvenience, invasiveness, lack of accessibility, and cost associated with CBCs, POC Hgb diagnostics have been developed to address some of these issues, but current systems all suffer from some combination of high-cost (handheld Hgb meters cost USD $30-$1000), inaccuracy, the need for blood sampling, and incompatibility with patient self-testing at home[5,7,24,25]. Therefore, a noninvasive, inexpensive, easily accessible (i.e. using only the native hardware of the

phone without the need for additional equipment) anemia diagnostic that accurately measures a patient's Hgb levels, especially as a patient self-test, can vastly improve the quality of life of chronic anemia patients. Furthermore, chronic anemia conditions such as sickle cell disease and β-thalassemia primarily affect underserved populations, further necessitating the need for such an inexpensive technology.

Towards that end, smartphone-based mobile health (mHealth) is steadily transforming the delivery of medical care worldwide, moving diagnosis of disease from the clinic to the home to potentially anywhere in the globe and facilitating enhanced patient-provider communication[26]. As examples, smartphone images of the skin, ears, and eyes have recently been used by physicians to remotely diagnose melanoma, ear infections, and corneal abrasions, respectively[10,11,27]. However, smartphone mHealth approaches have yet to non-invasively replace blood-based testing, which remains a major cornerstone of disease diagnosis in modern medicine. While the addition of specialized smartphone attachments and supplemental calibration tools may enable POC diagnosis and analysis of tissue and bodily fluid samples[12], the additional burden of blood and/or tissue sample collections combined with the additional cost and inconvenience associated with acquisition of the additional equipment prevents worldwide use of these potentially disruptive approaches. Therefore, the smartphone-based, patient-operated Hgb measurement system that I have developed, which requires nothing other than the native technology of the smartphone itself and is potentially capable of non-invasively replacing blood-based diagnostics, would transform the very nature of mHealth and delivery of healthcare worldwide. To that end, this technology will fundamentally shift the treatment paradigm and improve the quality of life of chronic anemia patients.

More broadly, by moving Hgb measurement from the clinic to the bedside, any patient in any location, at any time, now has access to an important health metric and may seek care accordingly. This technology has the potential to significantly impact individuals with chronic anemia, which disproportionately impacts underserved populations. For example, each year in the United States, approximately 2.25 million units of blood are transfused to 750,000 hematology/oncology patients, the majority of whom suffer from chronic anemia[28]. A Hgb measurement is taken prior to administration (typically using a CBC) of each blood transfusion to determine eligibility for a transfusion. Using this technology, these patients could monitor their anemia from the comfort of their own home, sending results directly to a physician to remotely assess their need for a transfusion, rather than through inconvenient, recurring, and costly clinic visits. In addition, patients with sickle cell disease, milder forms of thalassemia, and cancer patients undergoing chemotherapy, all of whom are already chronically anemic, may experience dangerous and precipitous drops in their Hgb levels from issues as benign as an upper respiratory viral infection and therefore require frequent Hgb monitoring[2]. The ability to self-test at home may save the patient from unnecessary clinic visits, as well as more quickly detect severe Hgb level drops. Moreover, this technology has far reaching global health applications as healthcare officials in low resource settings may use this technology to inform allocation of limited healthcare resources (e.g. transfusions, high-risk obstetrical services, etc.) for the patients with the most severe anemia. Furthermore, while this algorithm is currently focused on anemia, this approach can potentially be easily adapted to quantify other conditions that manifest in physical

exam findings and currently require blood tests (i.e. jaundice and cyanosis), demonstrating the versatility of this system and approach[29,30]

1.4 Innovation

Given the need for patient-operated, solely smartphone-based approaches that can non-invasively replace blood-based diagnostics, I have developed a smartphone app that quantitatively analyzes patient-sourced photos to enable a noninvasive, accurate, and accessible anemia diagnostic that requires only the patient's smartphone without the need for any additional equipment (i.e. smartphone attachments or calibration tools) or blood sampling. Here, a patient simply takes a picture of their fingernail beds, and the image analysis algorithm analyzes color data and image metadata to measure the corresponding Hgb level. This technology leverages the sophisticated imaging capabilities of existing smartphones to remotely diagnose and screen for diseases and encompasses several key innovations:

1.4.1 This algorithm "learns" and improves the accuracy of Hgb level measurements as more data is collected.

As images are obtained and matched with gold-standard level Hgb levels, as is necessary for personally calibrated algorithms, the dataset used to generate and "train" the Hgb measurement algorithm expands. Increasing the size of the image dataset ensures that a more representative sample of the general population is utilized to develop the algorithm, improving the overall accuracy of the proposed technique. Furthermore, as more images are added to the dataset, more sophisticated algorithms are applied to

facilitate this improvement. Thus, this innovative Hgb measurement method "learns" as the sample size increases.

1.4.2 Smartphone incorporation facilitates patient-operated Hgb measurement.

Current technology used to measure Hgb levels requires a trained healthcare professional to operate. Existing POC technologies, designed to bring Hgb measurement from a clinical laboratory directly to the patient, are designed for healthcare providers and none of these technologies are designed to be used at home by the patient as a self-test. The technology I have developed leverages the ubiquitous worldwide ownership of smartphones with the user friendliness and familiarity of smartphone apps to develop a system that allows patients to conduct sophisticated blood testing themselves, without requiring a costly, inconvenient, and potentially cost-prohibitive clinic visit. Bringing anemia screening and monitoring from the clinic directly to the patient has the potential to fundamentally shift the treatment paradigm for those suffering from, or those at-risk for, anemia.

1.4.3 Leveraging the sophisticated imaging capabilities of smartphones enables truly noninvasive Hgb measurement obviating the need for additional diagnostic equipment.

Physical examinations of clinical pallor have long been used as a qualitative metric for diagnosing anemia. Various studies conducted in low resource settings have attempted to semi-quantitatively assess anemia by correlating physician assessment of clinical pallor with presence of anemia (defined at various blood Hgb level cut-offs)[13-17]. The system described in this thesis utilizes robust multi-linear regression to correlate

color and image metadata from smartphone images of fingernail data to specific Hgb levels, representing the first case of quantitative Hgb level measurement using only images, let alone smartphone images.

1.4.4 Furthermore, personalized algorithm calibration allows for highly accurate monitoring of Hgb levels in those with chronic anemia conditions.

Imaging a single patient over time in conjunction with CBC Hgb levels enabled this technology to establish personalized correlations between that individual's fingernail bed color patterns and fluctuations and their Hgb levels. In this thesis, I present a method that will allow individuals suffering from chronic anemia to accurately monitor their anemia from the comfort of their home, potentially reducing costly clinic visits. This was accomplished by applying a personalized training scheme within algorithm development to improve accuracy.

1.4.5 This work enables Hgb measurement to be conducted on and by anyone, anywhere in the world, at any time, using only a smartphone.

The increasingly sophisticated imaging capabilities of smartphones coupled with their ever increasing worldwide ownership enables a powerful platform for image-based diagnostics. As no other software or hardware is needed, the system that I have developed obviates the need to visit a clinical setting to conduct testing. This fundamentally changes the treatment paradigm for individual's suffering from anemia, and serves as a starting point to address the treatment of other conditions which manifest in physical signs and can thus be potentially diagnosed by imaging alone. Overall, these points highlight the

novelty and significant technical innovation of this work, and demonstrate the improvement of this technology over traditional anemia testing.

CHAPTER 2. BACKGROUND AND LITERATURE REVIEW

2.1 Anemia

2.1.1 Anemia Biology

Anemia, characterized by low red blood cell (RBC) and blood Hgb concentrations, is the world's most common blood disorder, afflicting nearly one-third of the world's population, especially young children, the elderly, and women of childbearing age[1,31,32]. Due to numerous etiologies, including malnutrition, genetic hematologic diseases (e.g. sickle cell disease, beta-thalassemia), infections/inflammatory disorders (e.g., HIV, rheumatoid arthritis), cancer chemotherapy, and renal failure, anemia can become chronic and severe[2,19-22]. Left untreated, individuals suffering from anemia can suffer fatigue, weakness, cognitive disorders, and life threatening cardiovascular collapse in severe cases[18]. At-risk individuals, especially those with chronic conditions, therefore need to be regularly tested to prevent these symptoms and guide treatment strategies.

Causes of anemia can be broadly grouped into the following categories:

2.1.1.1 Impaired production of RBCs

A primary cause of anemia is the inability of the body to produce sufficient levels of hemoglobin[33]. The most common cause of this inability to produce sufficient levels of hemoglobin is iron deficiency[34,35]. As iron is a key component to the production of red blood cells (erythropoiesis), a deficiency of iron can lead to the inability of the body to produce functional RBCs and, thus, hemoglobin[36]. The primary cause of iron deficient

anemia is poverty, malnutrition, and famine[37]. In these cases, patients are not able to absorb sufficient iron from their diet, inhibiting erythropoiesis. These problems are common in low-resource settings, and are the primary cause for the significant correlation between anemia and low resource settings. Furthermore, these issues with iron deficiency can also be caused by vegetarian diets, issues with absorption, and chronic blood loss due to menstruation, issues that impact developed countries in addition to low resource settings[37]. Once discovered, iron deficient anemia can be easily and relatively cost-effectively treated via dietary or intravenous supplementation, even in low resource settings[38-41]. This highlights the importance of knowing ones hemoglobin level, as small dietary alterations can lead to significant improvements in quality of life.

Furthermore, anemia related to insufficient production of RBCs can be caused by multiple hematological diseases, some of which can be chronic and severe. One example of such a disease is beta thalassemia major, a blood condition caused by a mutation in the hemoglobin molecule, preventing its synthesis[2]. This leads to inefficient production of RBCs. Individuals with this disorder require regular blood transfusions to treat their disease. This allows patients to replace their insufficient RBC production with the transfused RBCs. These patients require frequent monitoring in order monitor their Hgb levels to ensure their treatment strategies are working properly and to ensure that their hemoglobin levels do not drop below clinically significant levels.

2.1.1.2 Increased Destruction of RBCs

While inefficient production of RBCs can be responsible for anemia, the opposite is also true. Increased destruction of RBCs can also lead to anemia. Destruction of RBCs,

known as hemolysis, can be caused by a variety of disorders. The two most common mechanisms of hemolysis are intravascular and extravascular hemolysis. In intravascular hemolysis, RBCs are destroyed within blood vessels due to mechanical trauma caused by interactions with the vessel wall, complement fixation, and infectious agents[42]. The more common method, extravascular hemolysis occurs when RBCs are marked for destruction and destroyed in the vasculature of the spleen[43]. Disorders that can lead to hemolytic anemia include; hemoglobinopathies such as sickle cell anemia; immune disorders such as transfusion reaction due to ABO-incompatible transfused RBCs; drug induced hemolysis caused by drug coating RBCs and provoking an immune response; and infection such as when the malaria parasite leads to the weakening of the RBC membrane[3,44-49]. In all cases, premature destruction of RBCs leads to lower levels of usable hemoglobin, as the hemoglobin can no longer functionally carry oxygen within the vasculature after it has been expelled from the RBC[50]. It is very important for patients to understand track their hemoglobin levels when hemolysis is confirmed, as some of these conditions may lead to sudden severe drops in Hgb level that can lead to life threatening complications

2.1.1.3 Bleeding

The final, and most intuitive, cause of anemia is simply bleeding [51]. In these cases, hemoglobin is simply being lost due to blood loss. Bleeding related anemia can be caused by either chronic or acute blood loss. Acute blood loss can occur for a variety of reasons, including injury, accidents, and surgery[52]. In these cases, it is important that patients monitor their hemoglobin levels after one of these events, as these events can lead to rapid, precipitous drops in hemoglobin that that can occur before a patient can get

to a clinic and have their Hgb levels measured. Chronic bleeding-related anemia can be caused by a variety of disorders, such as Von Willebrand disease, hemophilia, and gastrointestinal bleeding[53-55]. When bleeding is chronic, hemoglobin levels can gradually decrease overtime, and thus need to be monitored.

2.1.2 Anemia Monitoring

As a significant percentage of the world's population is anemic, and each of these affected individuals requires anemia testing/monitoring, a number of methods have been developed to measure Hgb levels.

2.1.2.1 Gold-standard anemia diagnosis

Table 1: Noninvasive Hgb measurement compares favorably with currently-used anemia monitoring technology.

Device	Noninvasive	External Device Free	Inexpensive	Accurate
Complete Blood Count[4]	**X**	**X**	**X**	✓
Hemocue[56]	**X**	**X**	**X**	✓
Anemocheck[6]	**X**	**X**	✓	✓
HemaApp[57]	✓	**X**	✓	**X**
Masimo Co-Oximetry[58]	✓	**X**	**X**	**X**
Conjunctival Analysis[24]	✓	**X**	✓	**X**
WHO Color Scale[59]		**X**	✓	**X**
Smartphone App	✓	✓	✓	✓

The current gold-standard diagnostic test for anemia is the complete blood count (CBC), conducted via a clinical hematology analyzer[4]. The clinical hematology analyzer is extremely cost prohibitive, especially in resource poor settings, where anemia is the most prevalent worldwide. In order to conduct a CBC on a clinical hematology analyzer,

one must first purchase the machine and rent facilities to house and power it. A trained laboratory technician must be employed to run each test, which consumes reagents that must be purchased. Additionally, a complete blood count requires a venous blood draw from the patient, which requires employment of a second trained technician to administer, and can cause discomfort in patients, especially younger patients[23].

2.1.2.2 Point of care anemia diagnosis

Due to the expense associated with the clinical hematology analyzer, point-of-care devices (i.e. devices that can be used at the time and place of patient care) have been developed to decrease costs and increase accessibility, particularly to resource-poor settings. These point-of-care devices, while more cost effective, bring about their own challenges. Photometry and colorimetry-based Hgb prediction methods, used in products such as the Hemocue and Anemocheck, typically require costly additional measurement tools to measure absorbance of various chemical reactions involving Hgb, in addition to requiring venous blood draws or finger sticks[5,6]. Hgb color scale methods such as the World Health Organization color scale are generally very cost-effective, but rely on human interpretation of color comparisons of a droplet of blood on paper to a printed color scale to estimate Hgb, leading to inaccurate Hgb estimations[60]. More sophisticated, non-invasive, spectrophotometric Hgb estimation techniques such as the Masimo Pulse Co-Oximeters, can accurately estimate Hgb, but these techniques are also cost-prohibitive[7]. Digital imaging based technologies, including techniques utilizing smartphones, have been explored as a simple alternative, but results have been insufficient to replace blood-based testing[25,57]. These technologies and their associated drawbacks are summarized in Figure 2 and Table 1.

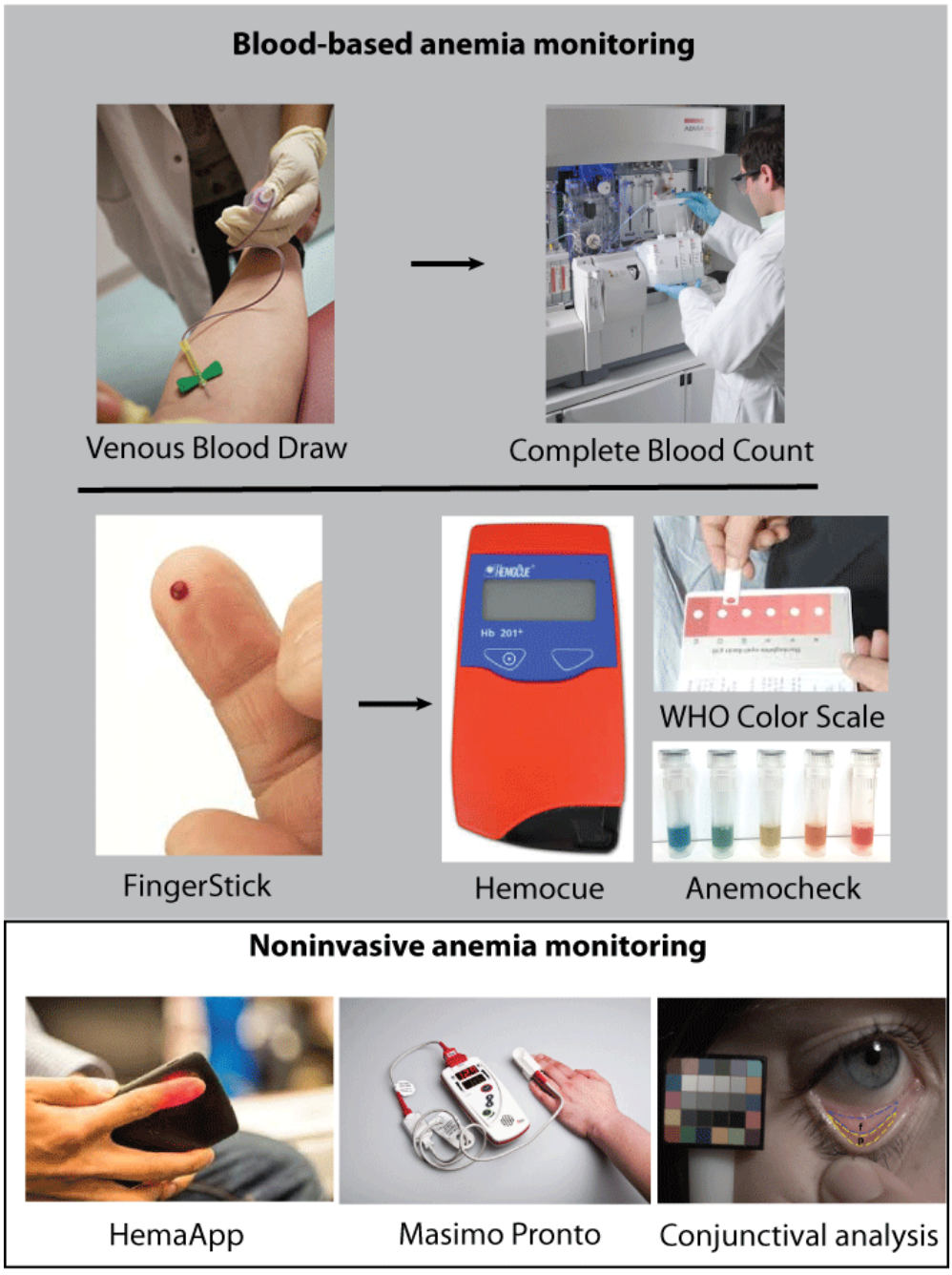

Figure 2: Current anemia monitoring technology. Traditionally, anemia is monitored via blood-based technologies in clinical settings. The gold standard anemia test is the complete blood count **(top row)**, which requires a venous blood draw. Other calorimetric technologies have been developed that require less blood and may be used in POC applications, such as the Hemocue, WHO color scale, and Anemocheck **(middle row)**. Other noninvasive technologies have also been developed that measure hemoglobin via light absorption or reflection off body regions, such as the HemaApp, Masimo Pronto, and conjunctival analysis in digital photographs **(bottom row)**.

2.1.2.3 Qualitative anemia diagnosis

While the gold-standard anemia diagnostic test is an extremely useful tool in determining a patient's treatment needs, physical examinations alone have long been used to aid in Hgb estimation. Multiple studies in low resource settings, where the gold standard anemia diagnostic tool is extremely cost prohibitive, have attempted to semi-quantitatively estimate Hgb in patients using results from physical examinations[13-17,61-65]. Skin color is determined by the 3 major pigments; melanin, by far the most impactful; carotenoids, the least impactful; and Hgb[66]. In specific regions of the body, such as the fingernail beds, conjunctiva, and palmar creases, melanocytes (cells that lie within the epidermis and produce melanin) lie dormant[67,68]. This causes the blood Hgb levels to have the greatest impact on skin tone in these regions. Pallor in these regions generally corresponds to a lack of Hgb in the blood vessels near the epidermal surface, providing a pseudo-anemia diagnosis. In the previously mentioned studies, pallor of the conjunctiva, fingernail beds, and palmar creases was assessed by physicians, who then characterized patients as "anemic" (defined by varying Hgb concentration ranges from study to study) or not. These characterizations were compared with Hgb levels measured using the gold standard. Quantitative data (sensitivity and specificity) were reported based on the accuracy of the physician's characterizations of anemia. While these findings were not able to accurately predict blood Hgb levels, they were successful enough to establish a correlation between clinical pallor and anemia and introduce physical examinations as a viable means to generally estimate Hgb.

2.2 mHealth diagnostic tools

The utility of physical examinations in estimating Hgb facilitates remote imaging as a means for anemia diagnosis. Mobile phone ownership is increasing worldwide. Currently, there are approximately 4 billion unique mobile phone subscribers worldwide, representing over half of the world's population[9]. In fact, there are more mobile phones in use than there are people worldwide, and half of those mobile phones are "smartphones". The fastest growing market for mobile phone subscribers is in low resource settings, where anemia is most prevalent. These smartphones incorporate sophisticated imaging software into their design, which can potentially be utilized by medical apps.

In fact, smartphones have been used in a variety of mHealth applications[69]. These apps can be classified into 3 major classes of devices or systems. 1) Apps that track medical record information and promote the maintenance of a healthy lifestyle, or facilitate patient-physician interaction and assist physicians with making a clinical decision; 2) apps that interface with an additional medical device and analyzes data or interprets a result; and 3) apps that use technology and computing power found on modern smartphones to turn the smartphone into a standalone medical device.

Apps that track medical data and promote healthy lifestyle maintenance are the most common mHealth apps. A classic example of an app that promotes maintaining a healthy lifestyle includes diet apps which track the user's body weight over time in order to assist the user in their goal of losing weight. These apps are also often used to gather and record sensitive medical data for the patient and physician. Examples of this include diabetes apps such as Daily Carb, which tracks key metrics for managing diabetes such as

meal times and readings for carbs, fiber, fat, and glucose which can then be shared with a physician[70]. Apps in this category can also serve to augment the patient-physician interface by simply offering the ability to search for doctors based on symptoms or triage diseases based on a list of symptoms, such as with the services ZocDoc and WebMD[71,72]. These apps can also be used to improve patient compliance to medication usage as well as refill prescriptions. These applications have presented a challenge to the medical device regulatory body of the United States of America, The Food and Drug Administration (FDA), as technology to assist with these issues has traditionally fallen into the realm of medical devices, and the market has become flooded with these technologies[73]. The FDA has responded by allowing exemptions to regulatory oversite for applications that pose no significant risk to patients should they fail and that fall into the categories of physician support or healthy lifestyle maintenance software.

Mobile app software that interfaces with external medical devices is another common mHealth modality. In fact, the majority of revenue generated by mHealth companies is driven by external device sales rather than in-app purchases[74]. A wide variety of technologies have been adapted to interface with smartphones, including such staple, sophisticated medical technologies such as echocardiograms, otoscopes, and pulse-oximeters. This enables direct collection, storage and transmission/communication of patient medical data directly to the patient's electronic medical records (EMR) as well as the physician. Furthermore, these systems potentially empower patients to control their own healthcare, as some of these systems may make available sophisticated medical tests that are not traditionally available in POC settings. These systems may also be used to enable new diagnostics and research tools. For example, microscopes that interface with

mobile phones enable analysis of patient samples remotely to look for conditions such as malaria. Also, microfluidic POC tests have been developed that interface with the camera on mobile phones to generate a readout, enabling remote and POC diagnosis of a variety of diseases.

Finally, the third classification of mobile medical apps is the least common, yet it forms the basis of the technology presented in this thesis. Very few technologies exist that can usejust the features and computing power of mobile phones themselves to turn the phones into a standalone medical device. The most prominent example of such a system is smartphone apps that are used to treat and monitor skin lesions in the context of skin cancer. In these cases, apps utilize the standard smartphone camera to monitor and characterize skin lesions based on a number of attributes in order to determine the likelihood that the skin lesion is malignant[75]. These systems represent major advantages over the other categories of mHealth apps. Unlike healthy lifestyle maintenance apps, they have the capability to accurately perform a diagnosis of a disease based on real-time, patient collected data that did not require a clinic visit or in some cases, even a trained medical professional. Unlike apps which interface with external medical equipment, these systems do not require that the user purchase an expensive piece of equipment in order to conduct testing. In this thesis, I have leveraged the imaging capabilities and the ubiquitous worldwide ownership of mobile phones with the diagnostic capability of physical examinations alone to develop a novel, non-invasive, easy-to-use, external device-free, smartphone application for diagnosing anemiawith the smartphone functions as a standalone medical device.

2.3 Image analysis in mobile medical devices

Traditionally, medical image analysis has been conducted by trained medical professionals using the human eye. The invention and improvement of sophisticated imaging modalities such as standard photography, magnetic resonance imaging (MRI), and computed tomography (CT), have enhanced the need for medical image analysis in the diagnosis and prevention of disease[76]. However, the large physical variability of various pathologies as well of the fallibility and potential fatigue of the human eye of medical experts necessitates automated image analysis via computation for more accurate diagnosis and classification of medical images. To that end, advances in central processing units (CPU) and graphics processing units (GPU) have enabled for complex, sophisticated computational image analysis, which has been able to pick up features in medical images that are missed by medical professionals. In some cases, knowing which features to look for is no longer a barrier to constructing a computational image analysis algorithm to extract features in medical images, as deep machine learning algorithms eliminate this need via automated algorithm training[77,78]. The only downside of these methods is that large datasets are typically required to validate that the algorithms truly work on all disease cases as opposed to overfitting to the limited cases they have been exposed to. Anemia represents an excellent case study in this field, as diagnosis of anemia may be conducted visually by looking for clinical pallor. Now, as camera technology has increased, image analysis of clinical pallor in medical images can be utilized to diagnose anemia from image. This method has led to the technology I present in this thesis in which sophisticated regression algorithms are incorporated into modern smartphones to diagnose anemia remotely using only a smartphone.

2.4 Patient populations requiring noninvasive anemia monitoring

Anemia afflicts a large percentage of the world's population, and instances of anemia are projected to increase. Anemia has many different causes that require different treatment approaches. Each distinctive cause of anemia can therefore lead to unique, innovative applications for this anemia diagnostic tool. Take, for example, resource-poor settings, where malnutrition is a major driver of anemia (Table 2)[1,79]. The lack of quality healthcare facilities in these low-resource settings makes field-accessible devices to perform analysis and diagnosis a necessity, indicating key beneficiaries of this technology[80].

Healthcare officials in low resource settings can use this Hgb level measurement technique as a means to inform the allocation of limited healthcare resources (Figure 3A). Healthcare workers will be sent out to areas where anemia is prevalent and screen the general population. This will allow patients suffering from anemia to be discovered who otherwise would not have been diagnosed due to lack of available resources, such as access to clinical hematology analyzers.

Table 1: Anemia affects a significant percentage of the world's population

Country	Percentage of population with anemia
USA	5.7%
United Kingdom	15.2%
Germany	12.3%
Russia	20.8%
Ukraine	27.3%
Indonesia	44.3%
Malaysia	38.3%
Nigeria	66.7%
Gambia	75.1%

Individuals at-risk for anemia in the general population can use this method as a general screening tool (Figure 3B). An at-risk individual could use this image analysis algorithm to screen themselves for anemia and seek medical treatment if necessary, rather than wasting time and incurring expense due to unnecessary hospital trips. As children

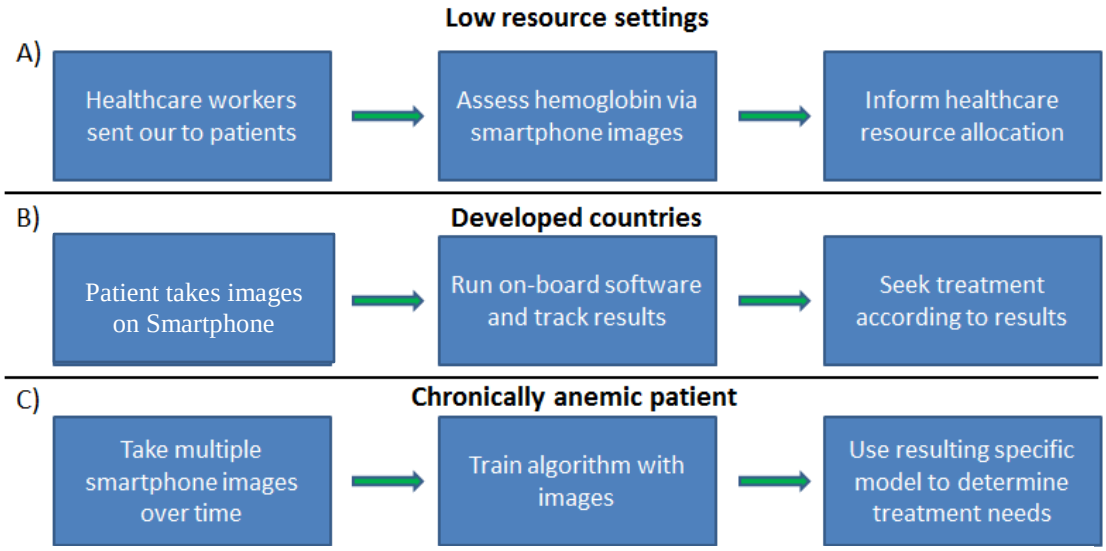

Figure 3: The non-invasive, image-based smartphone app for diagnosing anemia facilitates a wide range of use models. This Hgb estimation technique can potentially be used as a healthcare resource allocation tool in low resource settings **(A)**, a screening tool for anemia in general population at risk for anemia **(B)**, and as a monitoring tool for chronic anemia patients **(C).**

are all screened for anemia at around 2 years of age, an image-based, non-invasive test would allow parents to self-screen their children, eliminating unnecessary clinic visits and painful blood tests[81]. Chronic lead poisoning in children, an issue causing profound developmental delays and currently making national headlines, also leads to anemia. This algorithm could theoretically be used to monitor the health of children in areas with tainted water supplies[82].

Finally, In the case of a chronically anemic patient suffering from a hematologic disease in a developed nation, this tool could be used to self-monitor and allow the patient and physician to track the efficacy of drug and diet therapies (Figure 3C). A patient could potentially take multiple images, and compare the results with gold standard-measured Hgb levels to calibrate the smartphone app to their specific fingernail characteristics. This calibration would allow for much more accurate Hgb level estimation, which would permit an individual suffering from a disease casing anemia as well as the physician treating them, to accurately monitor disease progression and treatment efficacy remotely. The results of these tests could be used by the physician to remotely alter treatment regimens as necessary, without having to see the patient. A simple, non-invasive, inexpensive test for estimating Hgb levels has the potential to be used as a screening tool that has the capability to reach multiple user groups and improve the lives of patients suffering from a myriad of diseases.

CHAPTER 3. DETERMINE THE FEASIBILITY OF DEVELOPING AN IMAGE ANALYSIS ALGORITHM THAT CAN ACCURATELY PREDICT HGB VALUES FROM SMARTPHONE IMAGES OF SUBJECTS FINGERNAILS

3.1 Introduction

Physical examination alone has long been implicated in anemia diagnosis. Physicians have a long history of using physical examination of clinical pallor in a number of body regions including the conjunctiva, palmar creases, and the fingernail beds to qualitatively diagnose anemia[15,63,83]. These regions provide excellent anemia diagnostic capabilities becauset they do not contain melanocytes, the skin cells responsible for production of the pigment melanin and, thus, the skin tone of the individual[67]. As such, the primary contributor of skin color in these regions is Hgb (i.e. the molecule which gives blood its hallmark red color) of the blood flowing through the underlying vasculature. Thus, low Hgb levels lead to clinically observable pallor. Furthermore, as pallor is purely a color based phenomena, I and others have hypothesized that image analysis would be able to distinguish minute color variabilities related to Hgb fluctuations, such that Hgb level could be correlated with pallor[57]. Furthermore, I hypothesized that smartphone camera technology had become sufficiently advanced such that this correlation could be generated via smartphone images, leading to remote assessment of Hgb levels via a smartphone.

In this aim, I investigated the possibility of using the conjunctiva, fingernail beds, and palmar creases for assessing anemia via smartphone images, and determined the most appropriate body regions for conducting pallor-based Hgb analysis. In doing this, I took into consideration accuracy of results, ease of access to the body region, ability of an individual to self-image, and the computing power and user input necessary to identify the correct body regions of interest for analysis. Furthermore, in the process of identifying the optimal region of interest for correlating Hgb to pallor, I discovered a methodology for developing image analysis algorithms that can measure Hgb level based on color data from smartphone images of fingernails. In this chapter I report the reasoning behind choosing fingernail beds as the body region of choice for assessing anemia via smartphone images, the methodology for arriving at this choice, and the algorithm used to correlate fingernail color to Hgb level.

3.2 Materials and methods

3.2.1 Choosing a body region of interest

In order to choose the best region of interest on the body to be used to relate clinical pallor to Hgb level, images were taken of three body regions commonly used to diagnose anemia via physical examination. The regions of interest used were the fingernail beds, the conjunctiva, and the palmar creases. In order to take images of the fingernail beds, subjects were instructed to curl their fingers inward in order to control for blood flow alterations caused by hand position. Images of the conjunctiva were taken by requesting that subjects use their index finger to pull down their eyelids to expose the conjunctive. Images of the palmar creases were taken by asking subjects to flex their

hand outward, exposing their palms. All images were taken at a distance of approximately 0.5 m from the body region of interest. Red, green, and blue (RGB) pixel intensity color data was extracted from each regions of interest (ROI) within each image (corresponding to the body region of interest) and correlated with the gold standard-measured Hgb level of the subject. All images were taken with an Apple iPhone 5s (Apple, Cupertino, CA) using all default imaging settings.

3.2.2 Algorithm Development

In order to develop the Hgb estimation algorithm, average RGB pixel intensity data from ROIs within a single patient's fingernail beds were extracted from the smartphone images via a custom generated MATLAB (Mathworks, Natick, Massachusetts) function. This patient suffered from β-thalassemia, causing well defined fluctuations in Hgb level that could be used to correlate with pallor. Average pixel intensity values from the 3 color channels (RGB channels) were correlated with the gold-standard measured Hgb levels. I served as the initial patient in this proof of concept study, as my hemoglobin levels fluctuate rapidly and predictably due to a genetic blood condition that I am afflicted with (β-thalassemia major).

3.3 Results and Discussion

3.3.1 Fingernail beds are the optimal region of interest on the body for relating clinical pallor to presence of anemia.

In order to develop an image analysis algorithm that measures Hgb based on clinical pallor, an appropriate body region must be selected. This region must be known

to exhibit pallor when an individual is anemic, and must be easy to image. Three candidate body regions were chosen for this study, corresponding to an individual's

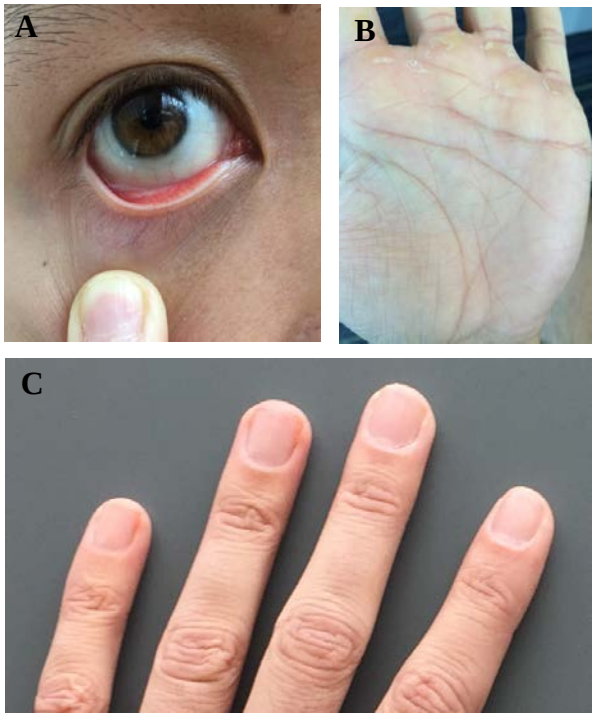

Figure 4: Images of various body regions associated with anemia. Images were taken of the **A)** conjunctiva, **B)** palmar creases, and **C)** fingernail beds in order to correlate pallor with Hgb level. **A-B)** These images highlight the issues with self-imaging of the conjunctiva, as well as the variability of the palmar creases. **C)** Furthermore, these images highlight the well-defined viewing window provided by the fingernail beds.

conjunctiva, palmar creases, and fingernail beds (Figure 4). The conjunctiva was the first body region studied, as it is commonly used to diagnose anemia via a physical examination. During imaging and data collection, it was difficult to collect RGB color data from the regions of interest for a number of reasons. First, it was difficult to explain to a subject exactly how far they need to pull their eyelid down in order to expose the conjunctiva. Furthermore, every individual tested pulled down their eyelid with different force, exposing different amounts of conjunctiva. This variability among individuals makes data extraction difficult as the region of interest changes drastically from person to

person. Also, looking towards implementation in a standalone smartphone app self-test, it is quite difficult for an individual to expose their conjunctiva and simultaneously take an image of it. For these reasons, the conjunctiva was ruled out as a region of interest for hemoglobin measurement. The next region of interest studied was the palmar crease. The palmar creases were difficult to analyze due to the fact that each individual had a different pattern of creases in their palms[84]. This would necessitate the development of a custom algorithm to identify each subject's palmar creases in order extract data from, a fact which immediately disqualified its use for anemia measurement. Finally, the fingernail beds were investigated as a region of interest for Hgb level measurement. While finger physiology varies slightly between individuals, the fingernails provide a consistent imaging window for the most part. These imaging windows are relatively large, facilitating simple RGB color data extraction from them. Furthermore, it would potentially be very easy for a user to use one hand to hold a smartphone and take an image of the fingernail beds on the other hand. For these reasons, the fingernail beds were ultimately chosen as the body region of interest for relating clinical pallor to anemia.

3.3.2 A Proof of concept study reveals that fingernail color is correlated to Hgb level in a chronically anemic and transfused β-thalassemia patient.

The proof of concept study for this project involved observing the fingernail bed color of a patient with a chronic hemolytic anemia disorder (β-thalassemia major) requiring RBC transfusion therapy. β-thalassemia is a chronic genetic disease caused by a mutation hindering or preventing synthesis of the β chain of Hgb, the blood's oxygen transporting molecule[2]. This inhibited Hgb synthesis leads to insufficient oxygen

transport in the patient's body. Treatment for this condition requires RBC transfusions to make up for the patient's insufficient Hgb production[2]. Over time, as the donated red blood cells are naturally destroyed by the body, the patient's blood Hgb levels drop. This drop continues until the patient's blood Hgb levels reach a clinically significant threshold (~10 g/dL), at which time the patient must receive another transfusion to avoid the manifestation of anemia symptoms. A patient suffering from β-thalassemia major

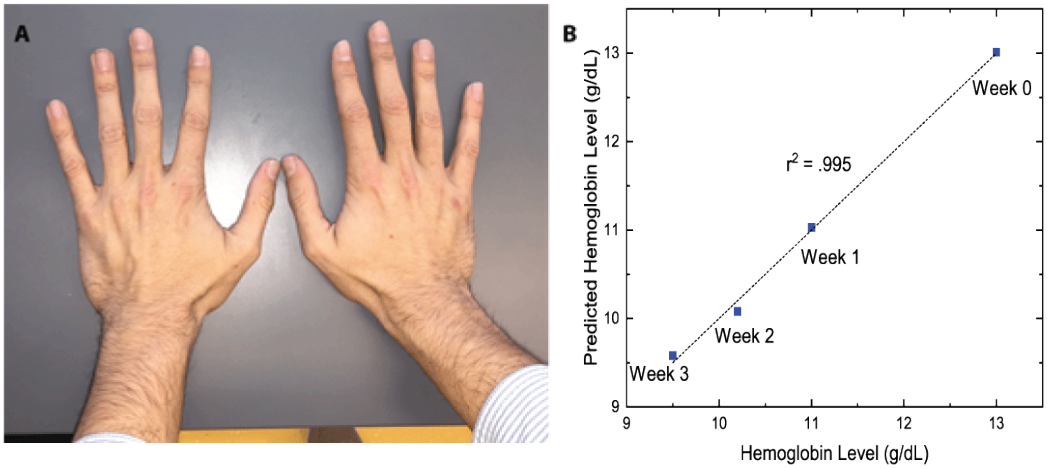

Figure 5: A non-invasive, image-based algorithm accurately predicts blood Hgb concentrations of chronically anemic patients. A) Representative patient image used to predict anemia. **B)** The predicted Hgb levels strongly correlate with Hgb levels determined by a clinical hematology analyzer (r^2 = .995) in a chronically anemic beta thalassemia patient over the course of several weeks after receiving a red blood cell transfusion.

represents an ideal test case for the hypothesis that Hgb levels can be accurately predicted using a strictly physical examination-based method due to the fact that their Hgb levels fluctuate very predictably through a wide concentration range over a relatively short period of time.

In this proof-of-concept study, images were taken on a smartphone of the patient's fingernail beds (Figure 5A) over the course of 3 weeks (corresponding to 1

transfusion cycle) and blood Hgb levels were estimated by correlating the "blue" channel pixel intensity to the patient's actual blood Hgb levels. These estimations were compared to blood Hgb levels measured via the gold standard blood test. Images and blood samples were taken 1 day post-transfusion (week 0), 1 week post-transfusion, 2 weeks post-transfusion, and 3 weeks post-transfusion (1 before the next transfusion cycle began). Hgb level estimates were able to be strongly correlated ($r^2 = 0.995$) with gold standard-measured Hgb levels (Figure 5B). The results from this study indicate that this technique can be used to successfully and accurately measure a single patient's hemoglobin levels. This proof-of-concept result served to legitimize my hypothesis that imaging fingernails alone is sufficient to measure blood Hgb concentration and opened the door of the future validation experiments presented in this thesis.

3.4 Conclusions and alternative approaches

Practical reasons presented in this chapter highlight the reasoning behind choosing the fingernail beds as the optimal region of interest on the body for relating clinical pallor to anemia. In addition to practical reasons relating to data extraction, the intended use of this technology was considered to exclude regions that are difficult to be self-imaged by an individual. Furthermore, quantitative data regarding the correlation between color data and gold standard-measured Hgb level presented in this proof-of-concept study showed that Hgb levels could be successfully estimated using only smartphone images taken of fingernail beds. Linear regression was used to relate blue pixel intensity value to gold-standard measured Hgb level. However, a major disadvantage of this proof-of-study is that the study was conducted on 1 subject. While I showed that Hgb levels could be accurately predicted, this study neglected many possible

confounding variables between different individuals that could affect Hgb estimation such as, imaging variability (e.g. imaging parameters such as the angle at which an image was taken, the distance from the subject, and the type of camera used to take the image), skin tone, background lighting, and nail size. These potential sources of variability are addressed in the following chapter, where a large scale clinical study was conducted to collect more data that encompassed a better representation of the anemic population who could potentially benefit from this tool. However, the data presented in this chapter proves that this is a viable technique for measuring Hgb levels.

CHAPTER 4. VALIDATE AND IMPROVE THE ANEMIA MONITORING ALGORITHM WITH A LARGE SCALE CLINICAL ASSESSMENT

4.1 Introduction

In order to validate the image analysis algorithm that relates fingernail color to Hgb level, more data needed to be collected from many different individuals in order to account for variability between individuals. Fingernail color can vary between individuals for a number of reasons, including thickness of the fingernail, abnormal discolorations (e.g. leukonychia, melanonychia, and bruising), and lighting conditions under which fingernail images are taken[85]. As proof of concept studies used to develop the algorithm relied on data collected from a single individual, parameters related to variability could not be included in the model. This caused a situation known as overfitting, where the model describes a given situation perfectly, but inadequately describes alternate situations. In this case, the image analysis algorithm developed in the proof of concept study was able to accurately measure Hgb in the single individual, but would be likely less successful if applied to a wider population exhibiting fingernail variability. In addition to being able to accurately account for variability within a population, increasing sample size enabled the usage of more sophisticated correlation algorithm, including multi-linear regression and machine learning techniques such as neural networks.

In this chapter, I detail a clinical validation study in which I increased the sample size of the dataset used to correlate fingernail color with Hgb. In addition to increasing the dataset, I discuss the techniques used to improve the algorithm to fit the larger sample size, and I discuss the ability of the algorithm to account for many common sources of variability in images of fingernail beds.

4.2 Materials and methods

4.2.1 Clinical Assessment of Noninvasive Technique for Estimating Hgb Level

A clinical assessment was conducted at Children's Healthcare of Atlanta, Emory University of School of Medicine, and Georgia Institute of Technology to relate fingernail bed color to Hgb levels. Patients with various anemia etiologies scheduled to have their Hgb levels measured via a CBC as part of their clinical care were recruited to this study (n = 265). Subjects were excluded by quality control measures if their images showed fingernail beds that were obscured or discolored due to leukonychia, nailbed injury, nail polish, darkening due to medication[86], etc. Exclusions were conducted to eliminate unnecessary variables that could obfuscate algorithm development. All CBC's were conducted using blood samples collected via venous blood draw. Verbal and written consent and assent were obtained from subjects and their parents (age permitting) in accordance with the Emory University Institutional Review Board (IRB) (00081226) and the Health Insurance Portability and Accountability Act (HIPAA) regulations prior to partaking in the study. After samples of patients' blood were collected to conduct their CBCs, two images were taken of those patients' fingernail beds. Smartphone pictures were obtained with the camera flash both on and off. All images were taken with an

Apple iPhone 5s using all default imaging settings. Prior to imaging, the auto-focus and brightness adjustment of the smartphone camera was activated by tapping the screen in the region of interest (i.e. the fingernail beds) in order to focus on the nailbed. To ensure

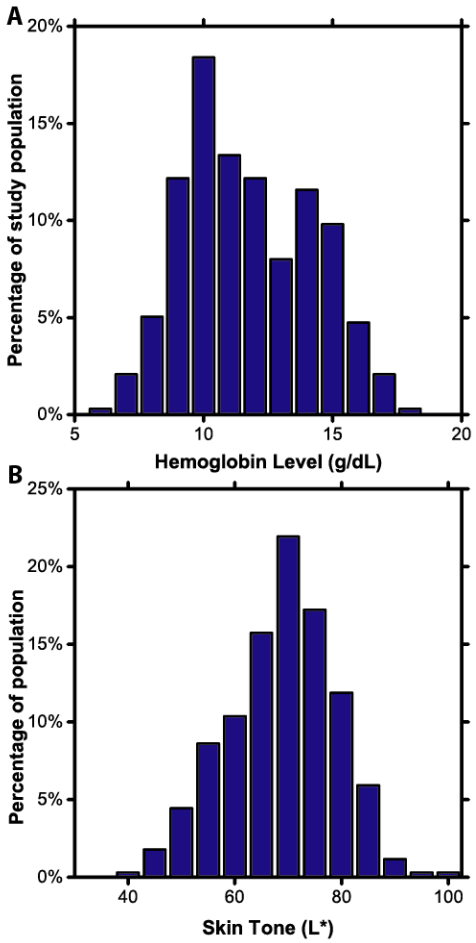

Figure 6: Hgb level and skin tone distribution of the study population. A) Subject's Hgb levels ranged between 5.9 g/dL and 16.8 g/dL. **B)** Subject skin tone was normally distributed. n = 337 subjects.

consistent images, each image was taken with the smartphone at a distance of approximately 0.5 m from the subjects' fingernail beds. Subjects were instructed to curl their fingers inwards with their palms facing upwards to control for possible alterations in blood flow caused by hand and finger positioning that could potentially affect the underlying color of the fingernail beds. Images were taken in clinic examination rooms,

where lighting conditions and room illuminants were relatively consistent. A digital light meter (Hisgadget, Union City, CA) was placed next to the subjects' fingernail beds to further ensure consistent background lighting conditions. Measurements taken from this external digital light meter were recorded to monitor background lighting conditions but were not incorporated into the Hgb level calculation. An additional 72 healthy subjects from Emory University and the Georgia Institute of Technology were tested using an identical protocol in order to ensure the Hgb level distribution of the study population better reflected the general population. CBCs were conducted on each subject prior to imaging and were analyzed via the same clinical hematology analyzer (Advia 2120i, Siemens, Berlin, Germany) used in the clinical study. All imaging was conducted in a room with similar lighting conditions to the clinic exam rooms, which was confirmed via digital light meter. Fingernail bed images and blood Hgb levels were analyzed in a total of 337 subjects. These subjects' blood Hgb levels ranged between 5.9 g/dL and 16.8 g/dL, (Figure 6A). Subjects' ages ranged between 1 and 60 years old. 167 female subjects and 170 male subjects were enrolled in this study. In 6 cases, fingernail polish was discovered after informed consent had been obtained, and these subjects were excluded from testing after study enrollment. In 1 case, an image labeled as having been taken with the camera flash on was discovered to have been taken with the flash off, resulting in this subject's data being excluded.

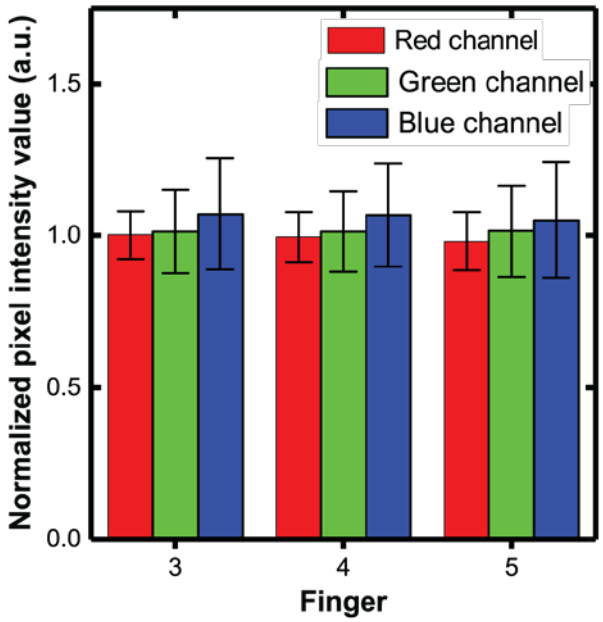

Figure 7: Color variability of fingernail beds is minimal across different fingers in the same individual. Color values in the red, green and blue channels were normalized to the second finger in each subject in order to compare color values from different subjects which were different due to variability within the study population. Fingers 3, 4, and 5 (the thumb was excluded from Hgb measurement) show little difference compared to finger 2. No statistically significant difference between color values across different fingers was found ($p > 0.30$ in all cases). Statistical significance was determined via two-tailed Student's t-test assuming unequal variance. $n = 10$.

4.2.2 Algorithm Development/Image Processing

Smartphone images were transferred or transmitted from the smartphone used in the study to a computer. Fingernail data, skin color data, and image metadata (i.e. data describing the camera settings at the time the image was taken) were extracted from fingernail bed smartphone images via MATLAB (Mathworks, Natick, MA). ROIs from which fingernail and skin color data were extracted were manually selected to ensure that fingernail irregularities were excluded from analysis. These ROI were selected from each finger excluding the thumb, and were 900 pixels^2, corresponding to approximately 10

mm^2 on the fingernail. The thumb was excluded due to the awkward hand position required to include the thumb in pictures. Color data was extracted from each region and averaged together across fingers for each subject. This was shown to be an acceptable method due to the low color variability between different fingers (Figure 7). An algorithm was then written in MATLAB utilizing robust multi-linear regression with a bisquare weighting algorithm to relate the image parameter data to CBC Hgb levels for each patient[87].

A uniform bias adjustment factor was also added to address the inherent variability in fingernail measurement. Two distinct use models and algorithms were applied for this Hgb measurement method: 1) a noninvasive, smartphone-based anemia screening test that does not require calibration with CBC Hgb levels, and 2) a noninvasive, smartphone-based, quantitative Hgb level diagnostic requiring calibration with CBC Hgb levels that enables chronic anemia patients to self-monitor their Hgb levels. Sampling strategies were used to generate the algorithm depending on the specific application.

4.2.2.1 Anemia Screening among the General Population

To develop the algorithm as a tool to screen for anemia, the entire study population (337 subjects) was randomly split into a "training" group (237 subjects) and a "testing" group (100 subjects). The training group was used to establish the relationship between image parameters and Hgb levels via robust multi-linear regression, much like the calibration phase of the personalized calibration study. A testing group of 100 subjects was used to validate the resultant algorithm. Validation was performed by

applying the smartphone algorithm to each testing image and comparing the algorithm generated Hgb result with the CBC Hgb result (i.e. determining the residual of the algorithm-based method). This process was iterated 1000 times with different randomly-selected without replacement training/testing groups to minimize residual error, thereby optimizing the parameters of the algorithm for anemia screening. Hgb measurements taken from the previously described personalized calibration study were not included in this anemia screening study.

4.2.2.2 Personalized Calibration of Smartphone Processing System

A personalized calibration approach was tested in two β-thalassemia major patients with chronic anemia currently undergoing chronic transfusion therapy, a healthy female subject with Hgb levels that fluctuated during her menstrual cycle, and a healthy male subject with consistent Hgb levels over an identical timeframe to assess the algorithm's capability to be accurately personalized and calibrated to that individual, regardless of their diagnosis or Hgb levels. Treatment for β-thalassemia major currently comprises of red blood cell transfusions to compensate for the patient's ineffective erythropoiesis[2]. Hgb levels in the chronic anemia patients fall throughout a 4 week transfusion cycle which was chosen as an appropriate time interval for this study. Prior to each imaging session, CBC Hgb levels were obtained from each subject via venipuncture. Color data and phone metadata were compiled, and a relationship between image data and CBC Hgb levels was established via robust multi-linear regression. This process was repeated for each individual using data from the 4 weeks of images to create a unique calibration curve personalized for that individual. This initial 4 week period is analogous to the "training" phase of the general population study, in which a personalized algorithm

is trained for each subject. Image parameter changes associated with Hgb level fluctuations specific to each person were related to perform algorithm calibration specific to each subject, thus improving the accuracy of Hgb level estimation. After the smartphone image analysis system was calibrated for each subject, Hgb levels were measured weekly over the next 4 weeks using the newly personalized algorithm. These Hgb level measurements were then compared to the CBC Hgb levels obtained at the same time to assess accuracy. This second 4 week interval corresponds to the testing phase of the general population study. This personalized calibration occurred over a total of 8 weeks.

4.2.3 Statistical Analysis

Statistical significance ($p < 0.05$) was determined via two-tailed Student's t-test assuming unequal variance. All statistical tests (calculation of regression correlation coefficients and Student's t-tests) were conducted using Origin Pro 2017 student version (OriginLab Corporation, Northampton, MA).

4.2.4 Study Approval

The clinical study regarding the application of the Hgb measurement algorithm was approved by the Emory University IRB (approval number 00081226). The study regarding the relation of finger temperature and exercise to measurement error was approved by the Georgia Institute of Technology IRB (approval number H17118)

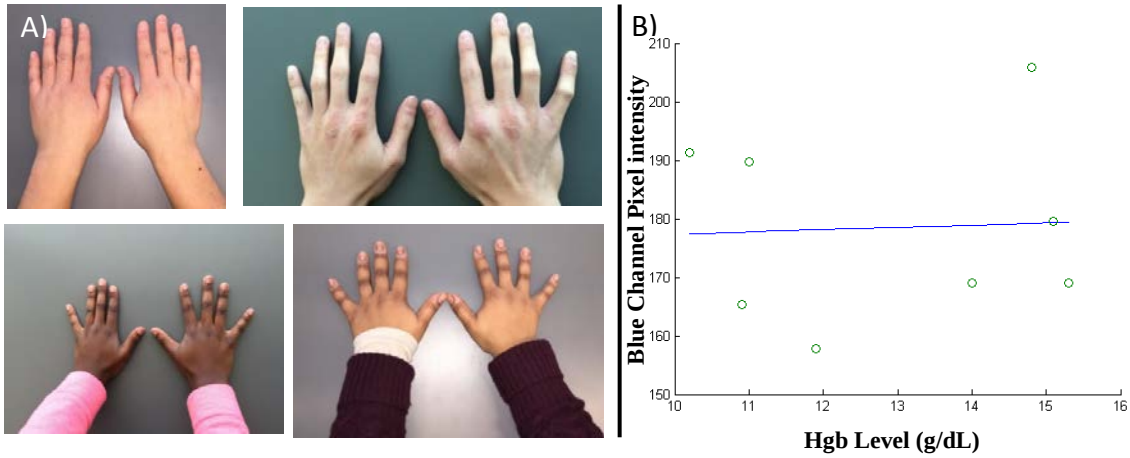

Figure 8: Variabilities within the tested population obscures Hgb level estimation. **A)** Images of the fingernail beds of a random sample of study subjects. Each image was taken using the same imaging conditions. However, lighting conditions and subjects' skin tones vary substantially between each image. **B)** When including a larger, more diverse sample population, blue channel pixel intensities show no correlation to goldstandard measured Hgb levels.

4.3 Results and Discussion

4.3.1 Variation between images of multiple individuals inhibits Hgb estimation based simply on one parameter.

The previously described method using linear correlation to relate blue channel pixel intensity to Hgb obtained in the β-thalassemia major patient study was applied to fingernail bed data from a group of healthy individuals (Figure 8A) in order to estimate their Hgb levels. These results did not show any significant correlation between blue pixel intensity and gold standard measured Hgb, indicating the presence of confounding variables within the experiment introduced by observing a larger population (Figure 8B). I hypothesized that these variables could potentially have arisen from the differences in skin tone between the individuals, differences in the fingernail physiology of each individual (such as fingernail thickness), and subtle differences in the lighting conditions

present at the time the images were taken. This prompted future experiments investigating the impact that skin tone has on Hgb level measurement, detailed in the next chapter.

4.3.2 Blood Hgb levels were accurately estimated in a group of healthy individuals and anemic patients at Children's Healthcare of Atlanta (CHOA) and the Georgia Institute of Technology.

Given my initial failure to correlate the average pixel intensity of a single color channel to blood Hgb levels in a larger, more diverse, population I decided to extract and explore other parameters from images of fingernails. To test this hypothesis that differences in skin tone and lighting conditions between diverse populations, I explored parameters related to these issues. In addition to looking at just RGB color data within the fingernail beds, I incorporated RGB color data from a control patch of skin located between the first and second knuckles on the middle finger.

This introduced skin tone as a variable in the model, allowing the algorithm to normalize for different skin tones. I believed that this a valuable model parameter due to the fact that different skin tones can relate to different pigmentation levels within fingernail beds. For example, melanocytes, which typically lie dormant within the fingernail beds, are more likely to actively produce melanin in individuals with darker skin tones[68]. In order to normalize for differences in lighting and imaging conditions between each image, metadata was extracted from each individual image file. Parameters extracted from the image metadata are used by the algorithm to correct for differences in external lighting conditions.

Adding additional parameters to the model increased the model's complexity. This increased complexity, while computationally expensive, had the potential to decrease the error (defined as the difference between the estimated Hgb level and the gold standard-measured Hgb level) in the Hgb measurement algorithm, as it encompassed more variables related to fingernail pallor in the general population. However, increasing the complexity of the model also increased the likelihood of overfitting the data, potentially leading to inaccurate Hgb level estimation. Overfitting occurs when ratio of model parameters relative to the number of observations is too high[88]. A perfect fit could potentially have been generated to the study subjects available. If applied to a larger population, however, the fit that worked well for the study size may not have accurately represented the larger population. In order to obtain a sufficient sample size of subjects to prevent overfitting due to the addition of multiple parameters to this image analysis algorithm[89], I conducted a large-scale clinical study at Children's Healthcare of Atlanta (CHOA).

4.3.3 Multiple linear regression was used to relate parameters extracted from clinical study-derived images of fingernail beds to blood Hgb levels. These relationships were used to develop a prediction algorithm for estimating anemia.

Multiple linear regression was used to model the relationship between the previously described parameters (independent variables) and blood Hgb concentrations (dependent variable)[90]. The ordinary least squares method was used to estimate this relationship[91]. This method finds coefficients that, when multiplied by each unique parameter and linearly combined, minimize the difference between the estimated variable and the observed variable determined via gold standard Hgb estimation (1). A bisquare

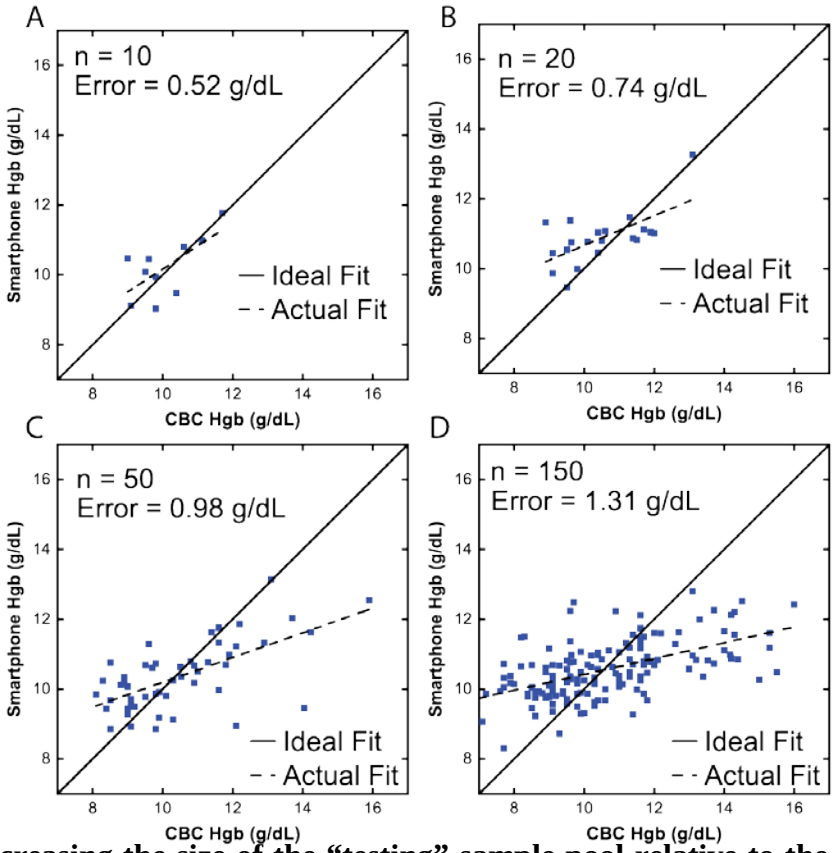

Figure 9: Increasing the size of the "testing" sample pool relative to the "training" pool increases error and bias in the prediction algorithm. A) 10, B) 20, C) 50, and D) 150 images were randomly selected and held aside from images of n = 273 subjects taken with the camera flash on. The prediction algorithm was trained with the remaining samples and the fit was tested on the held aside sample images. This process was repeated 1000 times and the best-case-scenario is shown. Average error of 1000 trials = ~1.5 g/dL in all cases. Solid line represents ideal prediction of predicted Hgb level = measured Hgb level. Dashed line indicated linear fit of the data. Decrease in slope of linear fit from ideal fit indicates experimental bias.

weighting algorithm was used to dampen the effects outliers have on the model, using a technique known as robust linear regression (2)[87]. Testing this method required splitting up the subject pool into randomized "training" and "testing" subject groups. The data from the training subject pool was used to model the relationship between the image parameters and Hgb concentrations as previously described. Once the relationship was established and a prediction algorithm determined, the relationship was applied to the testing subject group to attempt to predict their Hgb levels without using the gold

standard data. The data presented here shows the predicted Hgb values of the testing group plotted against the gold standard-estimated values (Figure 9). This process was repeated 1000 times to ensure that luck did not play a factor in "guessing" random images that happened to correlate better than others due to inadvertent experimental variability. In the best-case scenario, the prediction algorithm can estimate the blood Hgb concentration of 10 patients to within 0.52 g/dL (Figure 9A). Increasing the size of the training subject pool relative to the testing subject pool leads to an increase in the error of the best-case scenario observation, leaving the accuracy of the average of 1000 random iterations nearly identical at ~1.5 g/dL (Figure 9A-D). This indicates that the clinical study was of a sufficient sample size due to the fact that increasing the number of trained observations had little impact on the overall accuracy. As I increased the size of the testing pool relative to the training pool, the error in the best case scenario increased. However, with a sample size of n = 150, representing >50% of the total number of subjects, it was possible for the algorithm to predict Hgb levels to within 1.3 g/dL, a level of accuracy approaching that of similar, more expensive devices (Figure 9D)[92]. In addition to observing an increase in the error of the best-case-scenario as size of the testing pool relative to the training pool is increased, I observed a decreased slope in the linear fit. The practical implication of this observation is that subjects who have very high or very low Hgb levels are being predicted closer to the population's mean Hgb level than they should be. This indicates the presence of bias in the experiment.

$$HGB_{Prediction} = \sum_{i=1}^{n} C_i * P_i \quad (1)$$

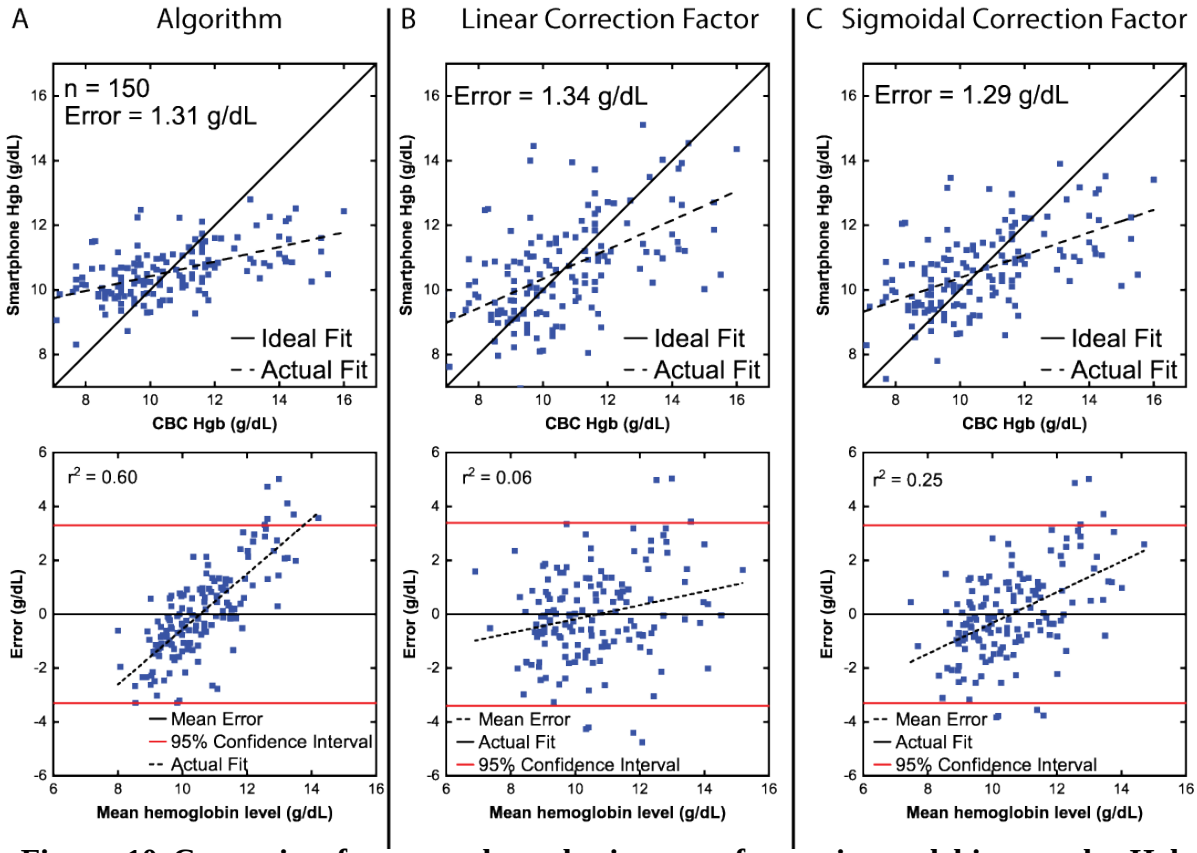

Figure 10 Correction factors reduce the impact of experimental bias on the Hgb prediction algorithm. A - Top) Predicted Hgb levels of 150 subjects were reported. A-**Bottom)** The slope of the linear fit of the data points in the Bland-Altman plot **(bottom)** is positive, indicating presence of experimental bias. **B-C Top)** Correction algorithms do not significantly alter error in the prediction algorithm. **B-C Bottom)** Correction algorithms reduce slope in the bland Altman plots, indicating mitigation of experimental bias. **Top)** Solid lines indicate ideal fit of measured Hgb level equals predicted Hgb level. Dashed lines indicate actual linear fit. **Bottom)** Hgb level difference is defined as the difference between predicted Hgb level and measured Hgb level of a subject. Mean Hgb level is the average of the predicted and measured Hgb levels for a subject. Solid black lines indicate 95% limits of agreement (1.96 x Standard deviation of the Hgb difference). Solid red lines indicate mean Hgb level difference. Dashed lines indicate actual linear fit.

Where n is the number of parameters in the model, $P_i \ldots P_n$ represents each parameter in the model, and $C_{i,\ldots}$ C_n represents the coefficient relating each parameter to the gold standard Hgb estimation.

$$weight = (\,abs(r) < 1) * (1 - r^2)^2$$

$$Where\ r = \frac{error}{tune * s * \sqrt{1-h}} \tag{2}$$

Where *error* is the difference between the predicted Hgb level and the gold standard estimated Hgb level, *h* represents leverage values from a least-squares fit, *tune* is the default tuning constant, and *s* is the standard deviation of the error.

4.3.4 Transforming the prediction algorithm data partially corrects for the experimental bias without sacrificing accuracy

In order to address the observation that subjects with Hgb levels near the boundary of the population's Hgb range tend to be predicted closer to the mean (i.e. predicted too high near the low end of the range and too low near the high end of the range), I applied transformations to the original Hgb measurement data (Figure 10A) in an attempt to accurately predict Hgb levels across the entire range of Hgb levels. A linear (3) (Figure 10B) and sigmoidal (4) (Figure 10C) correction algorithm based on the predicted Hgb level distance from the population mean was applied to the results from the prediction algorithm on 150 subjects. Both correction factors reduced the bias, as shown by the reduction of slope in the Bland – Altman plots (a plot primarily used to illustrate compare the performance of a new method to the gold standard), while preserving the error (Figure 10Bottom)[93]. While this represented an improvement in the prediction algorithm, this correction did little to address the cause of the bias in the data.

$$HGB_{Corrected} = HGB_{Prediction} + (HGB_{Prediction} - HGB_{Mean}) \quad (3)$$

$$HGB_{Corrected} = HGB_{Prediction} + \left(\frac{n}{1 + e^{-x}}\right) \quad (4)$$

Where n = max value of transformation, and $x = HGB_{Prediction} - HGB_{Measured}$

4.3.5 Altering the testing and training protocol to account for Hgb level distribution decreases error and reduces experimental bias.

I hypothesized that the experimental bias leading to large error at the boundary of the study population's Hgb level range could be due to the Hgb level distribution in the study population. Due to the nature of their conditions requiring treatment at CHOA, the majority of the study subjects were mildly anemic. Training populations were randomly selected from the study population to develop the prediction algorithm; therefore, the prediction algorithm was trained based on data that had a clear anemic majority. I believed that the least squares algorithm was being solved in a way that fits the predictions to the training data set. In other words, Hgb measurement data was flawed due to the flawed Hgb level distribution of the training data. I amended the protocol for selecting training samples from the study population to investigate this hypothesis. The study population was broken up into 3 groups based on gold standard-measured Hgb level (<9 g/dL, 9 – 12 g/dL, and >12 g/dL). An equal number of subjects from each group was randomly chosen to train the algorithm using the previously described method. This ensured a constant Hgb level distribution among the training population. The resulting prediction algorithm was applied to the remaining subjects. This training method was

successful, resulting in an average error of 0.89 g/dL in 100 patients and a reduction in bias (Figure 11).

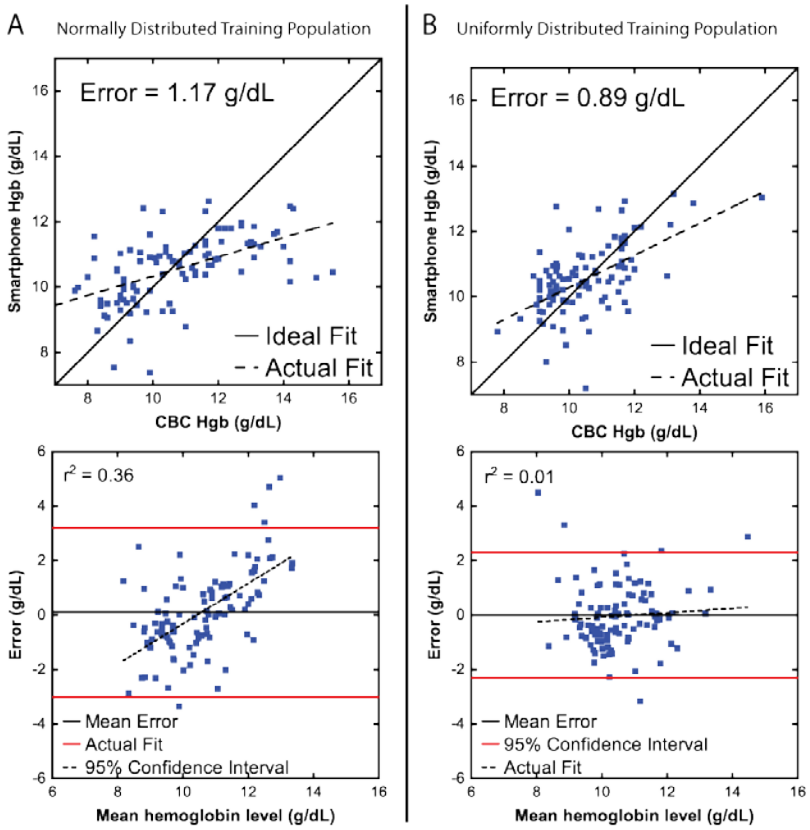

Figure 11: Training population Hgb level distribution affects accuracy of the Hgb prediction algorithm. A) Hgb level predictions of 100 subjects tested using a prediction algorithm trained with a subject group possessing normally distributed Hgb levels **(Top)**. The slope in the Bland-Altman plot **(Bottom)** indicates experimental bias. **B)** Hgb level prediction of 100 subjects tested using a prediction algorithm trained with a subject group possessing uniformly distributed Hgb levels **(Top)**. When trained with a population possessing a uniform Hgb level distribution, average error decreases **(Top)**, and bias is reduced **(Bottom)** relative to a normally distributed training population. **Top)** Solid lines indicate ideal fit of measured Hgb level equals predicted Hgb level. Dashed lines indicate actual linear fit. **Bottom)** Hgb level difference is defined as the difference between measured Hgb level and predicted Hgb level of a subject. Mean Hgb level is the average of the predicted and measured Hgb levels for a subject. Solid black lines indicate 95% limits of agreement (1.96 x Standard deviation of the Hgb difference). Solid red lines indicate mean Hgb level difference. Dashed lines indicate actual linear fit. N = 100.

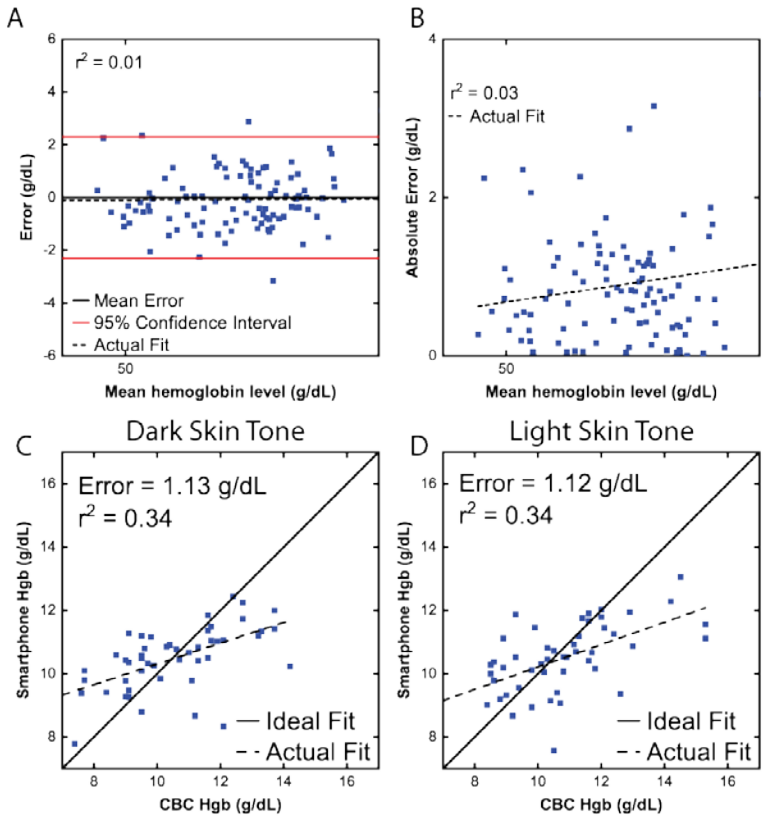

Figure 12: Skin tone has negligible impact on Hgb estimation error. A) Difference and **B)** absolute difference of Hgb levels show a weakly positive correlation to skin tone (shown on a 0 to 100 scale with 0 being absolute dark skin tone and 100 being absolute light skin tone. Hgb estimation algorithms trained with subjects possessing **C)** dark or **D)** light skin tones show minor difference in error. Dashed lines indicate linear fit. Solid lines indicate ideal fit of predicted Hgb level equal to measured Hgb value.

4.3.6 Skin tone has little impact on error in Hgb level estimation

I hypothesized that another potential cause of the observed experimental bias stems from the fact that individuals with darker skin tones are more likely to have abnormally pigmented fingernail beds[68]. I employed automated analysis of skin tone to account for this. The CIELAB color space was designed as a device independent color space capable of approximating human vision[94]. Much like the human brain, this color space distinguishes between light and dark ("L" value), between red and green ("A"

value), and between blue and yellow ("B" value), by using a transformation on the RGB values in each image based on the illumination source of the image. The lightness, or "L" value, most accurately represents human perception of skin color and serves as an ideal variable linearly modeling skin tones in a population[95]. Using these non-biased skin tone values, I investigated whether a subject's skin tone had any correlation with the error associated with the Hgb level prediction for that subject. I found that the "L" value of a subject's control patch of skin, in which a higher value correlates to lighter skin tone, had little impact on the absolute Hgb level difference with a weakly positive correlation between lighter skin and greater Hgb level and absolute Hgb level difference (Figure 12A-B). I also amended the training group selection procedure once more to determine if grouping subjects by skin tone could lead to more accurate predictions algorithms based on skin tones. Subjects with light skin tones and dark skin tones were grouped together and the previously described training/testing procedure was conducted on each group. Here I show that the error associated with the algorithm trained with individuals with dark skin tones was nearly identical to the error associated with the algorithm trained with individuals with light skin tones (Figure 12C-D). These results indicate that melanocyte activity and subsequent pigmentation of the fingernail beds and, thus, skin tone did not play a significant role in the observed experimental bias in this Hgb level estimation procedure.

4.3.7 Machine learning techniques do not improve Hgb level measurement accuracy given the current sample size of the study population.

While the accuracy of Hgb measurement that I have reported is within clinically acceptable levels for a POC device, more sophisticated regression methods were

investigated in an attempt to improve Hgb measurement. In order to assess the viability of more sophisticated Hgb measurement models, Hgb measurement was conducted using 2 methods of machine learning that are capable of determining unanticipated non-linear relationships between fingernail data and Hgb levels. Neural Network regression utilizes a network of hidden nodes that can approximate nonlinear relationships between a set of input parameters[96]. Using a training set of inputs (fingernail bed image parameter data) and outputs (gold standard measured Hgb levels), a "learning" algorithm utilizing Levenberg – Marquardt backpropagation and implemented in MATLAB was used to model the relationship between fingernail bed image parameters and gold standard-measured Hgb levels. This relationship was tested on a group of 100 images to estimate their Hgb levels, similar to the training and testing procedure outlined previously in this aim using linear regression. Using this method, the Hgb estimation error increases to 1.6g/dL as compared to robust multi-linear regression (Figure 13A-B). Bagged decision tree ensemble regression utilizes multiple decision trees corresponding to different subsets of features (fingernail bed parameters in this case) trained on small random groups of the overall training set (i.e. the image pool), a technique known as bootstrap aggregation, and averaging the regression results of the different trees in order to reduce overfitting[97]. When this method was applied to a testing image group, the Hgb estimation error increased to 1.26g/dL compared to the robust multi-linear regression case (Figure 13A,C). This indicates that neither method has improved upon the success of the simple robust linear regression model. However, the Hgb level distribution of the training group in all three cases was normally distributed, which I have shown to be less effective than a uniformly distributed Hgb level distribution.

4.3.8 Anemia Screening Using the Smartphone Image-based Algorithm when additional subjects are included to normalize Hgb level distribution

This system has the capacity to serve as a noninvasive anemia self-screening tool for use by the general population or at risk populations. With a single smartphone image and no personalized calibration step, smartphone Hgb levels were measured to within ±2.6 g/dL with a bias of 0.0 g/dL of CBC Hgb levels in 100 patients mixed with healthy subjects (Figure 14A, r = 0.79; Figure 16, Table 3), defined as the 95% limits of agreement. This noninvasive approach represents a greater degree of accuracy than reported accuracy levels of existing invasive anemia screening methods used in low resource settings such as the World Health Organization (WHO) color scale[60].

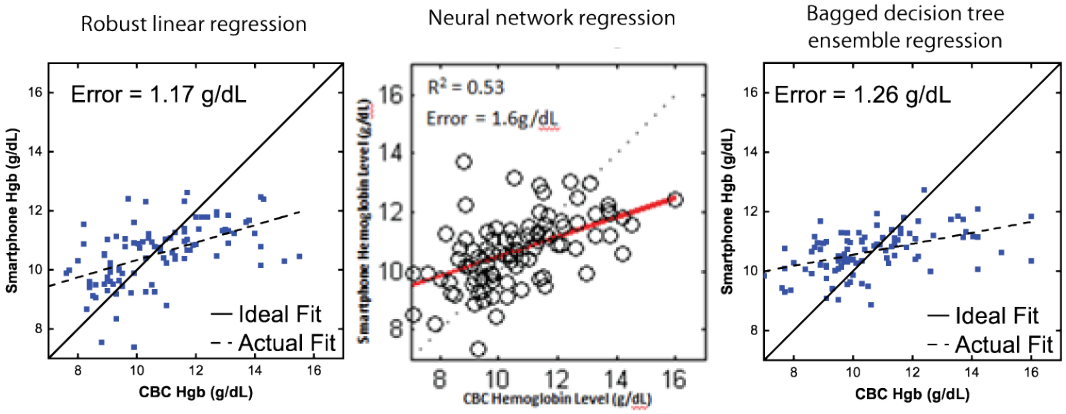

Figure 13: More sophisticate methods of regression seem to have little impact on Hgb estimation accuracy. Hgb estimation conducted via **(B)** neural network regression (Error = 1.6g/dL) or **(C)** random forest ensemble regression (Error = 1.26g/dL) fail to improve upon Hgb estimation using **(A)** robust multi linear regression (Error = 1.17g/dL). n = 100 subjects.

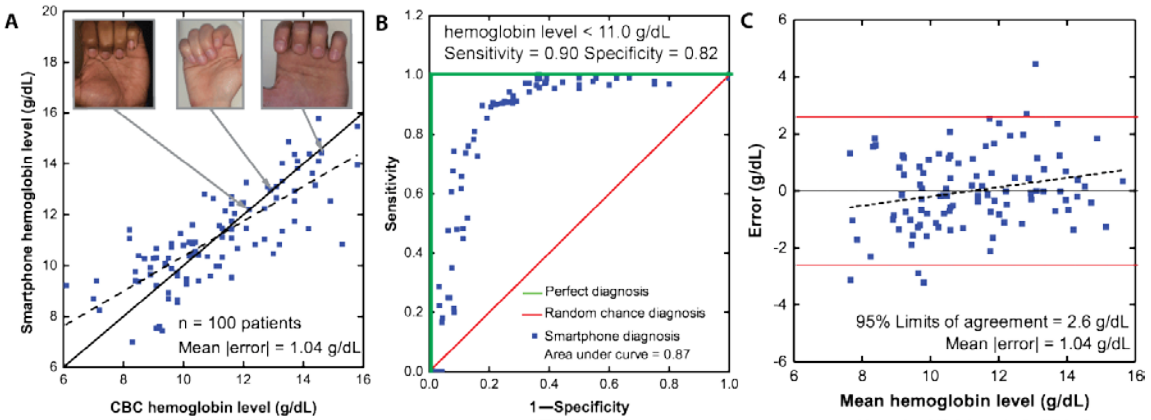

Figure 14: The smartphone-based image analysis algorithm accurately measures Hgb levels. A) The smartphone image analysis algorithm measures blood Hgb levels to within ±1 g/dL of the CBC Hgb level (r = 0.79). The solid line represents the ideal result where smartphone Hgb level is equal to the CBC Hgb level whereas the dashed line represents the actual data fit. Inset images illustrate example patient-sourced photos that were used to calculate Hgb level measurements. **B)** The Receiver Operating Characteristic (ROC) analysis graphically illustrates the algorithm's diagnostic performance against a random chance diagnosis (red line), with an area under the curve of 0.5, and a perfect diagnostic (green lines), with an area under the curve of 1. In the case of this noninvasive smartphone app Hgb measurement system (black line), the area under the curve of 0.87 suggests viable diagnostic performance of this algorithm. When using the WHO Hgb level cutoff of < 12.5 g/dL, the sensitivity of the test is 94% (95% CI, 89% – 100%), n = 100 patients. **C)** Bland-Altman analysis reveals minimal experimental bias with zero average error, indicating that Hgb measurement is not biased in any direction. The dashed line represents the relationship between the residual and the average of Hgb level measurements obtained from the CBC and the algorithm (r = 0.22). The solid red lines represent 95% limits of agreement (2.6 g/dL). n = 100

Table 3: Diagnosis profile of clinical assessment subjects. This study enrolled 337 individuals with a wide variety of diagnoses. Study subjects enrolled consisted of 162 patients with Sickle Cell disease (e.g. type SS, SC, Beta+, Beta 0), 34 patients various other anemias (e.g. Beta Thalassemia major and minor, Microcytic, macrocytic, normocytic, aplastic, iron deficient, and general anemia), 54 instances of several malignancies (e.g. Leukemia, Lymphoma, Sarcoma, Neuroblastoma, Germ Cell Tumor), as well as 15 patients suffering from various other blood conditions (e.g. Deep Vein Thrombosis, hyperbilirubinemia, Hypogammaglobulinemia, idiopathic thrombocytopenic purpura, Purpura fulminans, Pulmonary embolism, Neutropenia, Spherocytosis, Thrombocytopenia, Von Willebrand disease). Additionally, 72 healthy control subjects were enrolled in the study to ensure that a wide range of Hgb levels were represented.

Disorder	Number of Patients
Hematologic Diseases	**211**
Sickle cell disease	*162*
Hgb SS	126
Hgb SC	21
Hgb Sβ+	8
Hgb Sβ0	5
Beta thalassemia	*17*
Major	12
Minor	5
Thrombocytopenia	6
Aplastic anemia	2
Deep vein thrombosis	2
Hemophilia	2
Microcytic anemia	2
Neutropenia	3
Pulmonary embolism	2
Anemia (unspecified)	1
Diamond blackfan anemia	1
Hemolytic anemia due to immunosuppression	1
Hyperbilirubinemia	1
Hypogammaglobulinemia	1
Iron deficient anemia	1
Macrocytic anemia	1
Normocytic anemia	1
Pancytopenia	1
Paroxysmal nocturnal Hgburia (PNH)	1
Purpura fulminans	1
Spherocytosis	1
Von Willebrand disease	

Table 3 continued…	
Oncologic Diseases	**54**
Acute lymphoblastic leukemia	29
Hodgkin's lymphoma	6
Acute myeloid leukemia	5
Diffuse large B cell lymphoma	4
Neuroblastoma	3
Osteosarcoma	2
Anaplastic ALK-positive large cell lymphoma	1
Burkitt's lymphoma	1
Chronic myeloid leukemia	1
Ewing's sarcoma	2
Extragonadal germ cell tumor of mediastinum	1
Germ cell neoplasm of the left testicle	1
Hepatoblastoma	1
Lymphoma (unspecified)	1
Rhabdomyosarcoma	1
Sacrococcygeal germ cell tumor	1
Spindle cell sarcoma	1
Synovial cell sarcoma	1
T lymphoblastic lymphoma	
Wilm's tumor	
Healthy Control	**72**

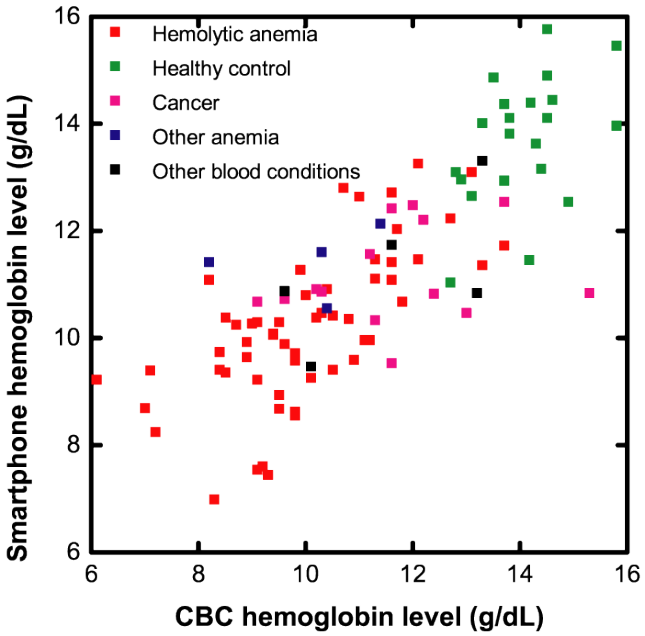

Figure 16: Diagnosis profile of Hgb measurement subjects. Subjects with hemolytic anemia, healthy controls, cancer, other anemia (e.g. aplastic anemia), as well as various other blood disorders (e.g. such as thrombocytopenia, deep vein thrombosis, and hemophilia) participated in the study. These data represent the diagnosis profiles of the subjects shown in Fig. 2. n = 100.

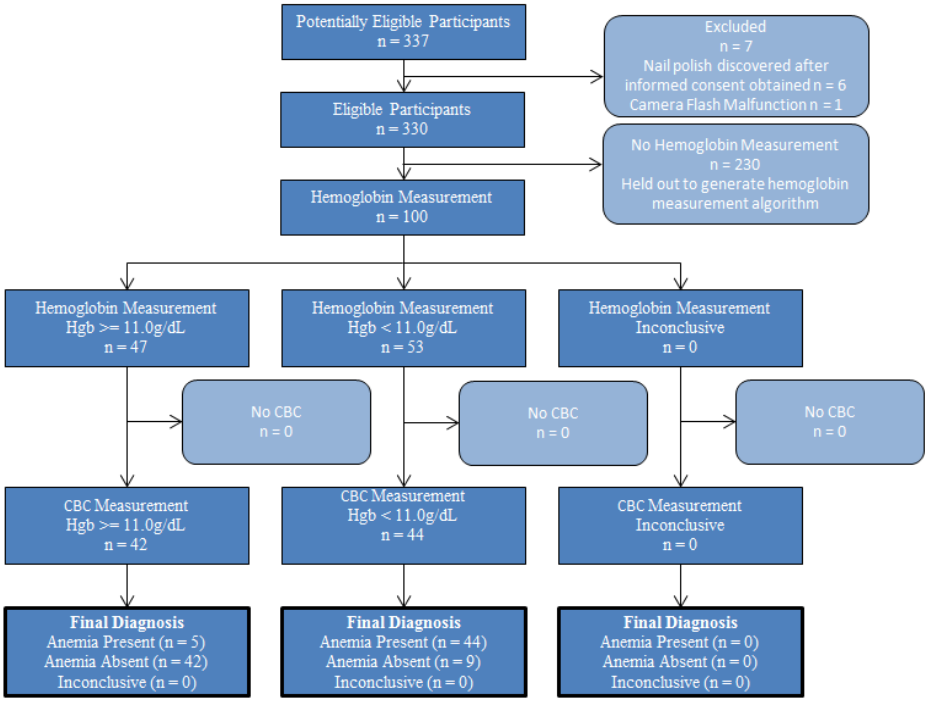

Figure 15: STARD diagram for the general population study when defining a positive test for anemia as < 11.0 g/dL

Moreover, receiver operating characteristic analysis revealed that this test

achieves a strong diagnostic performance with an area under the curve of 0.87 (Figure 14B) and highlights the accuracy of this technology throughout the entire range of tested Hgb levels. Additionally, there was minimal correlation between patient Hgb levels and smartphone-measured residual (r = 0.23), indicating that the algorithm performance remained consistent throughout range of tested Hgb levels (Figure 14C). Notably, when using a cutoff of <11.0 g/dL to define anemia, a well-established Hgb level threshold (Figure 14B, Figure 15),[16] the sensitivity and specificity of the system to detect anemia was 90% (95% CI, 82% – 98%) and 82% (95% CI, 71% – 93%), respectively. Using the WHO Hgb level cutoff for anemia of 12.5 g/dL, the sensitivity of the test improves to 94% (95% CI, 89% – 100%), indicating the potential for this test to serve as a noninvasive screening tool for anemia (Figure 17)[98].

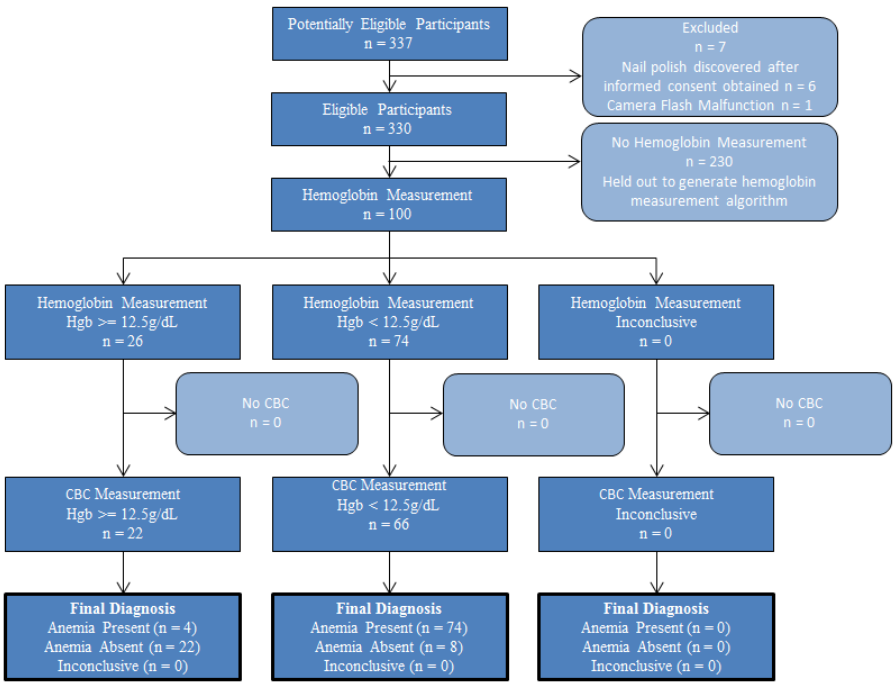

Figure 17: STARD diagram for the general population study when defining a positive test for anemia as < 12.5 g/dL

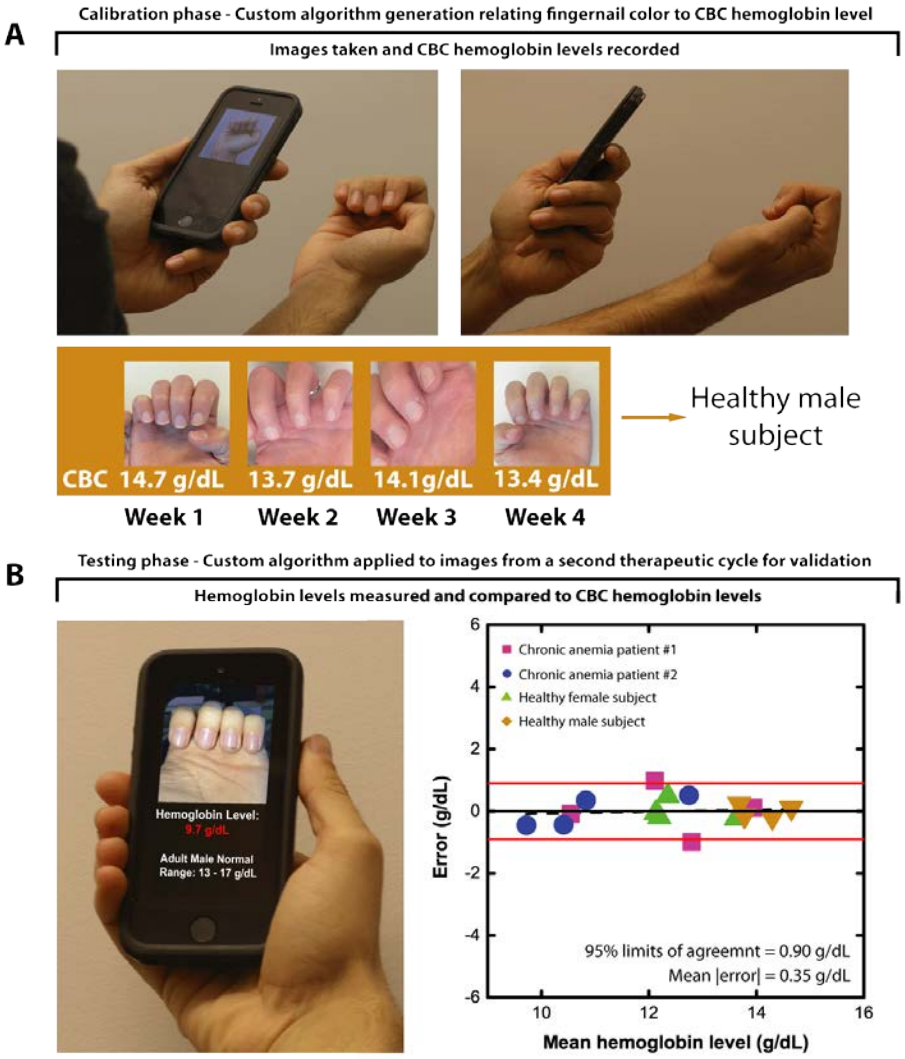

Figure 18: Adding a personalized calibration step to generate a patient-specific algorithm further improves the accuracy of Hgb levels measurement and is ideal for chronic anemia patients. A) Healthy and chronically transfused anemic patients were monitored over four weeks (i.e. over the course of a therapeutic blood transfusion cycle). CBC Hgb levels (white text) were used in conjunction with the images to generate a personalized algorithm for each individual. **B)** The patient-specific algorithms were used to measure Hgb levels over a subsequent blood transfusion cycle. This patient-specific calibration improved the average error of Hgb level measurements to within 0.35 g/dL of the CBC Hgb level. Bland-Altman analysis shows negligible experimental bias in the data. The zero average error (solid black line) indicates the Hgb measurement of the smartphone app is not biased towards any direction. The dashed line represents the correlation (r = 0.02) between the residual error and the average of Hgb level measurements obtained from the CBC and the algorithm. The solid red lines represent 95% limits of agreement (0.9 g/dL). n = 4 patients, 4 measurements per patient.

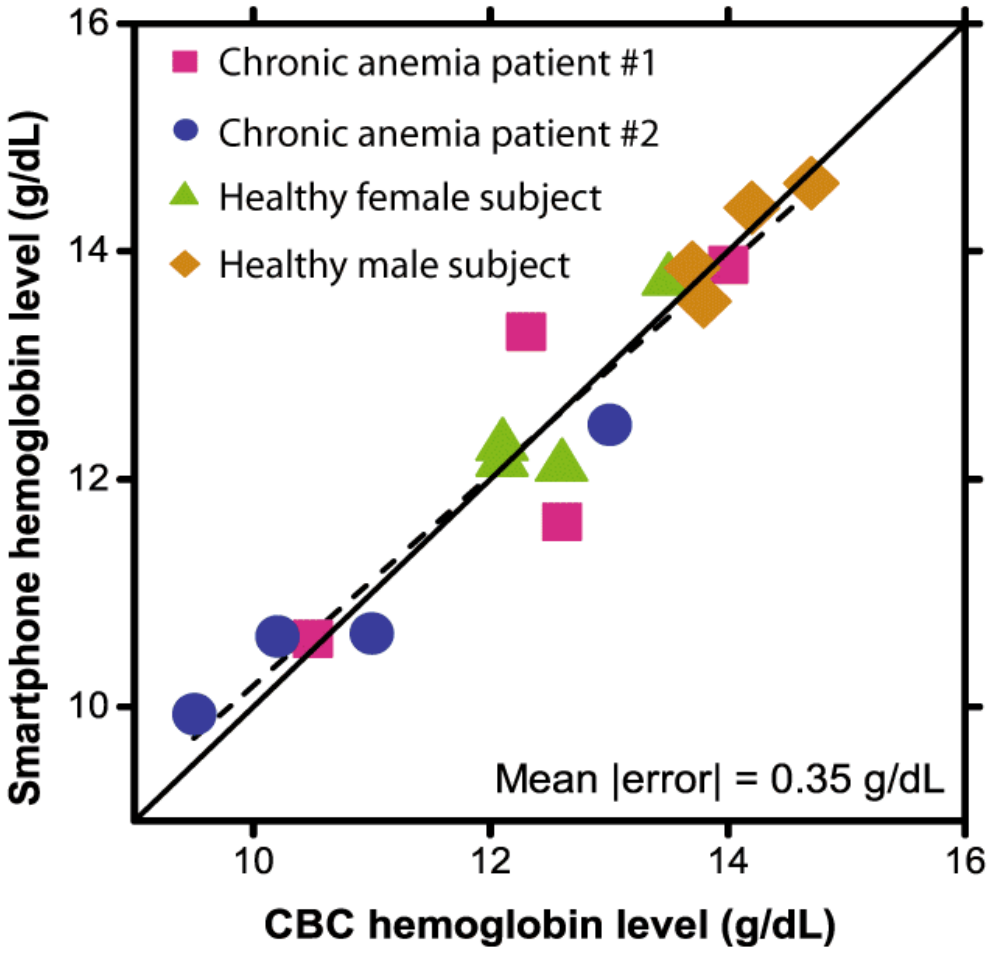

Figure 19: Patient-specific Hgb level measurements are in agreement with CBC measurements in the study population throughout a physiologic range of Hgb levels. This patient-specific calibration improved the accuracy of Hgb level measurements to within 0.35 g/dL of the CBC Hgb level (r = 0.95). The solid line represents the ideal result where smartphone Hgb level is equal to the CBC Hgb level. The dashed line represents the actual data fit. n = 4 patients, 4 measurements per patient.

4.3.9 Personalized Hgb Level Measurements Using the Smartphone Image-based Algorithm

Furthermore, the smartphone-based algorithm was calibrated for each subject (4 subjects total) over the course of 4 weeks and achieved personalized and accurate Hgb level measurements, enabling long-term serial monitoring of Hgb levels (Fig. 4A, Figure 18, Figure 19). Overall, when used in this manner, this system achieved a level of

accuracy of ± 0.90 g/dL with a bias of 0.0 g/dL compared to CBC Hgb levels (Figure 20), again, defined by the 95% limits of agreement (i.e. the Hgb level difference from the gold standard that 95% of smartphone measurements will fall between), representing an improvement on the reported accuracy of current invasive, point-of-care Hgb tests, such as Hemocue[7], and clinically-used noninvasive methods such as the Masimo Radical 7[92].

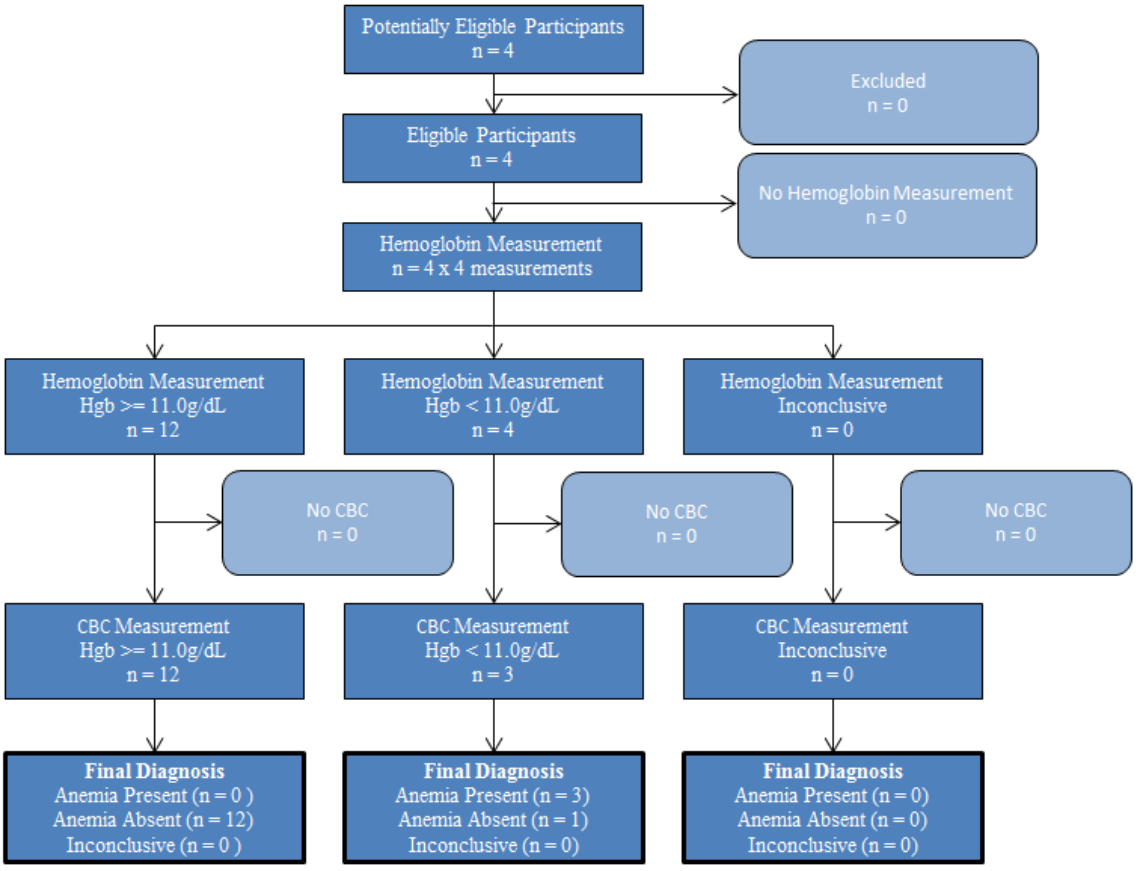

Figure 20: STARD Diagram for the personalized calibration study when defining a positive test for anemia as < 11.0 g/dL

Additionally, the smartphone Hgb level measurement residual did not correlate with the average between each patient's CBC Hgb level and smartphone Hgb level with ($r = 0.02$), indicating that residuals were not biased for any specific range of Hgb levels

(i.e. algorithm performance remained fairly constant throughout the entire physiologic range of tested Hgb levels). Furthermore, this degree of accuracy falls below a clinically significant threshold of ±1 g/dL,[60,99,100] suggesting that this system can potentially be considered interchangeable with the CBC Hgb level given further study[101]. Furthermore, 88% of measurements fall within Clinical Laboratory Improvement Amendment (CLIA) allowable total error of ±7%[102], indicating that upon further refinement and completion of additional testing, this technology may potentially be viable for at-home and clinical use.

4.4 Conclusions and alternative approaches

There are possible potential sources of error introduced by the design of this clinical study. Given the medical histories of the patients visiting the hematology/oncology clinic, many of the subjects in the clinical study had low blood Hgb levels. If the majority of patients used to train the algorithm had low blood Hgb levels, this could potentially cause the algorithm to artificially lower blood Hgb estimates of future healthy patients. Also, due to regional demographics, the majority of patients enrolled in the study was African-American and possessed relatively dark skin tones. Additionally, the majority of these patients with dark skin tones were suffering from diseases whose symptoms include low blood Hgb levels. Since these patients were used to "train" the algorithm, it is possible that future patients with dark skin tones will have artificially low measured Hgb levels. Finally, this experimental design relies on the smartphone's ability to normalize for the different imaging conditions mentioned previously (e.g. automatic white balancing, auto-exposure, auto-focus), rather than using a color normalization feature located within the image (e.g. a color/whiteness standard). However, these automatic normalization functions are not perfect, which could

potentially lead to small variations in fingernail color and skin tone unrelated to the blood Hgb concentration. Additionally, some of the participants in the clinical study were on therapeutic drugs that have been shown to affect fingernail color. While subjects were typically excluded if they had nailbed discolorations or abnormalities, it is possible that a subject was included that had a very subtle discoloration due to medications.

A possible source of the experimental bias I have identified is the distribution of Hgb levels in the study population. I have attempted to engineer the training data set such that Hgb levels in the training population are uniformly distributed, in order to ensure the prediction algorithm is not biased towards any Hgb level. However, because so many subjects were used in the training population, the left-over testing population still had a Hgb level distribution heavily influenced by anemic Hgb levels. In order to ensure that both the training and testing subject groups have uniform Hgb distributions, as well as enough subjects to prove the algorithms utility as a Hgb screening tool, I will begin seeking out healthy subjects with Hgb levels in the normal range to take part in future iterations of this study. The previously mentioned techniques will be used to generate new prediction algorithms based on the new study populations. It is possible that this method of Hgb estimation will not achieve the accuracy necessary to be a clinically valuable diagnostic tool. This could occur either due to the inability to account for experimental bias or that pallor variability between patients is simply too great for this method of Hgb estimation to account for. If this is the case, I will investigate more sophisticated modeling methods to generate the prediction algorithm in greater depth. One such modeling technique involves the use of machine learning and neural networks, as discussed previously, which takes advantage of the ability of computers to iteratively

learn more complicated models and extract features and parameters within images of fingernail beds that are not readily apparent to the human eye[103]. While these techniques did not improve Hgb measurement accuracy when applied to the study population presented in this chapter, it is possible that the more sophisticated, machine learning-based Hgb estimation methods will improve when trained with a much larger patient sample size encompassing a uniformly distributed Hgb level distribution. These techniques have shown great promise in diagnosing a variety of medical problems from diabetic neuropathy to skin cancer[77,78]. By inserting a regression layer on top of the deep neural network, covariates (such as age, anemia etiology, camera type, etc.) can easily be factored in to the regression to optimize the algorithm. Although it is traditionally expected that thousands of patients are needed for deep learning to bear fruit, researchers have seen improvements over state of the art algorithm just using hundreds of patients[104]. Therefore, the use of more subjects representing numerous etiologies, as well as healthy controls (50% male/50% female) will ensure the scientific rigor of this approach. Moreover, multiple recordings per patient, which has not been implemented in the current testing protocol, have the capability to enhance such algorithms, while cross validation ensures that I will not over-train on a given population. This will be attempted once additional subjects with healthy Hgb levels are enrolled in the study and the study population contains a more uniformly-distributed Hgb level distribution. I will utilize the vast amount of computing power at my disposal at the Georgia Institute of Technology and Emory University to develop these complicated models to solve the regression problem this project presents. Overall, these data highlight the ability of clinical pallor to be correlated quantitatively with blood Hgb levels in subject's who are both healthy and

suffer from a wide range of disorders. These results indicate that this method may be used to noninvasively measure Hgb and diagnose anemia and that future studies may improve Hgb measurement accuracy such that this technology may one day replace blood-based testing.

CHAPTER 5. PERFORM USABILITY TESTING AND INCORPORATE THIS ALGORITHM INTO A TRANSLATIONAL, EASY TO USE, AND INEXPENSIVE SMARTPHONE APPLICATION.

5.1 Introduction

In order to achieve its potential as a noninvasive diagnostic tool capable of remotely screening for anemia, this image analysis algorithm must be incorporated into a medical device that patients and healthcare providers can use. Smartphone ownership is increasing worldwide, and many medical devices have been designed that are incorporated into smartphones, or interface with mobile technology. As this image analysis algorithm for measuring Hgb levels utilizes only smartphone images of a subject's fingernail beds without the use of any external equipment, incorporation into a mobile app was a natural fit to allow for noninvasive self-diagnosis of anemia. The ability to function using a smartphone alone represents a substantial benefit to a user as they already have the medical device (the smartphone) in many cases and must only download an app to unlock this additional capability of their smartphone.

In this chapter, I discuss the integration of this image analysis algorithm into a fully functioning standalone smartphone app as well as the diagnostic accuracy of this app compared to the accuracy of a physical examination by a physician. I also discuss the user interface considerations that went into app development, as well as the incorporation of the app in multiple operating systems. I conclude with a discussion of the challenges

that come along with using multiple smartphone models that contain different camera hardware and the challenges of accounting for this hardware variability.

5.2 Materials and Methods

5.2.1 App Development

5.2.1.1 Software

The Hgb level measurement algorithm was incorporated into mobile apps. The open source integrated development environment (IDE) Android Studio (Google, Mountain View, CA) was used to develop a beta version of the Hgb measurement app in the Android operating system. The proprietary IDE Xcode (Apple, Cupertino, CA) was used to develop a beta version of the app in the iOS operating system.

5.2.1.2 Beta Testing

Design consideration that ultimately led to the process flow described below were determined by assessing the functional and technical requirements of a Hgb measurement app, as well as convenience and usability requirements determined via beta testing. Beta testing was conducted on a group of individuals who were comfortable with use of a smartphone but had no knowledge of application or software development. Users were given basic instruction on how to use the app and asked to provide feedback regarding usability of the app. This process was used to generate iterative improvements to the process flow of the app. The primary outcomes of this preliminary beta testfocused on how information was displayed and how instructions were presented to the user. In response to beta testing feedback, each set of instructions presented in the app process

flow is now presented to the user in a large, easy-to-see text box on the smartphone display. In some cases, images are included with these instructions that visually demonstrate the actions a user must take. For instance, the hand position that was used in the clinical study is demonstrated to the user to facilitate mimicking of the correct hand position to obtain accurate Hgb measurement.

5.2.1.3 Process Flow

The process flow for the apps that I have developed on both iOS and Android platforms is as follows and can be summarized by Figure 21:

1. The user is prompted to enter a patient ID, which allows patient data to be stored and collected for future analysis based on a deidentified number. This number prevents the subject's personal information such as name, date-of-birth, sex, etc. to be linked to data collection in any way.
2. The user is prompted to take an image with the camera flash on. The smartphone's default camera app is opened which allows the user to take an image, view the image, and choose whether or not to use said image. The image is then transferred back to the app where the user yet again has the option to choose retake the image.
3. The user is then prompted to manually choose ROI within the image. The user must then tap the image on each ROI (i.e. the patients 4 fingernails in the image excluding the thumb). This will trigger green boxes to be drawn around each region of interest.

4. The user will then be prompted to check and make sure the image looks is adequate for Hgb measurement. The user must check to make sure that fingernails are present in the image, fingernail beds are clearly visible (e.g. no glare/flash reflection, imaging artifacts, etc.), and the boxes are completely contained within the fingernail beds. If the image is determined to be unsatisfactory, the user has the option to retake the image or to redraw the boxes and reselect the regions of interest corresponding to the fingernail beds.
5. The user is prompted to measure the Hgb level of the fingernail image. This data can be recorded by the user.

After conducting the Hgb measurement using the app, the app stores the data for future analysis. The color data, image metadata, and the Hgb result are stored in a text file on the internal storage of the smartphone. The image taken by the user is also saved on the internal storage of the smartphone for future reference.

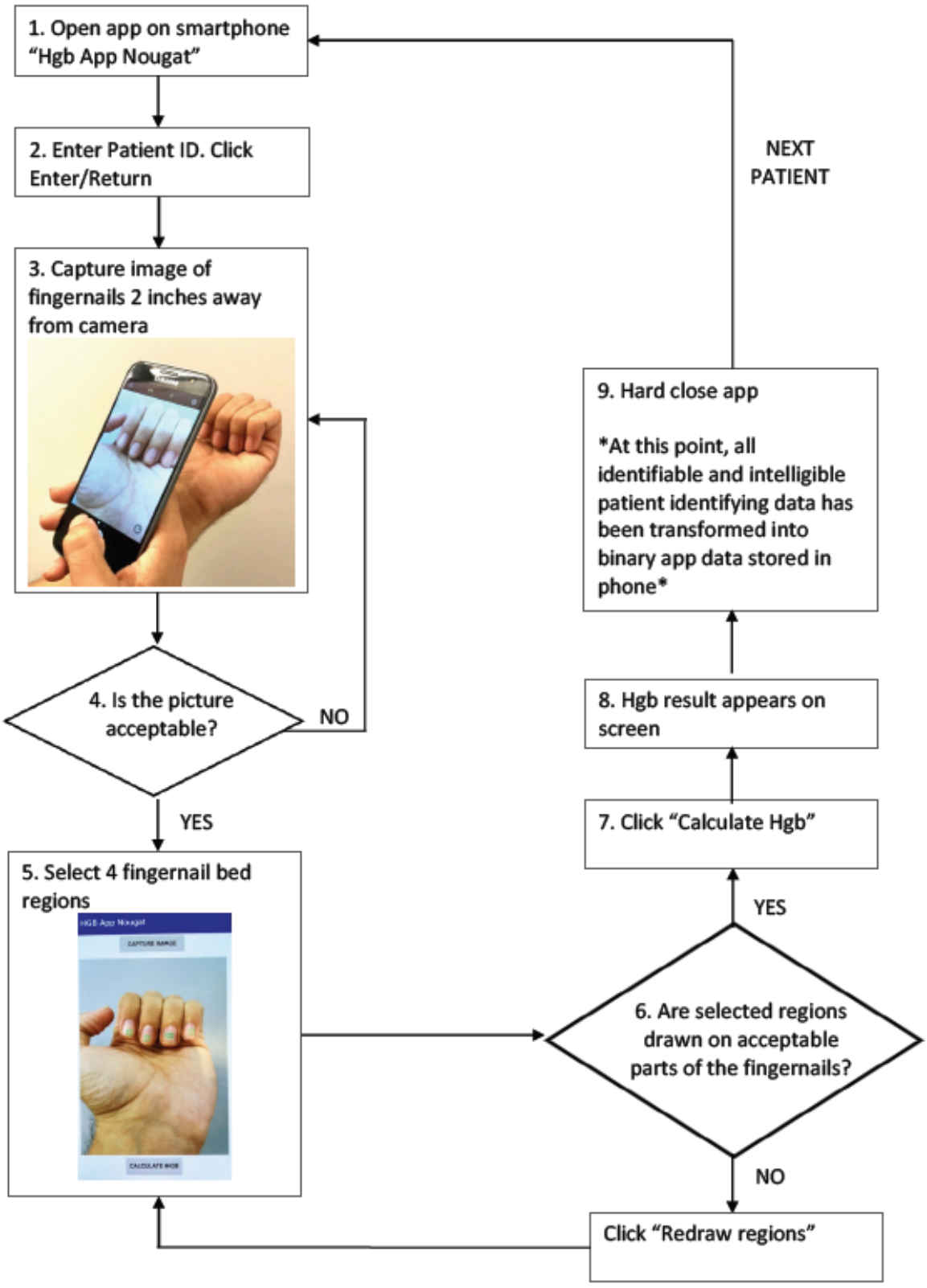

Figure 21: Smartphone app for diagnosing anemia process flow

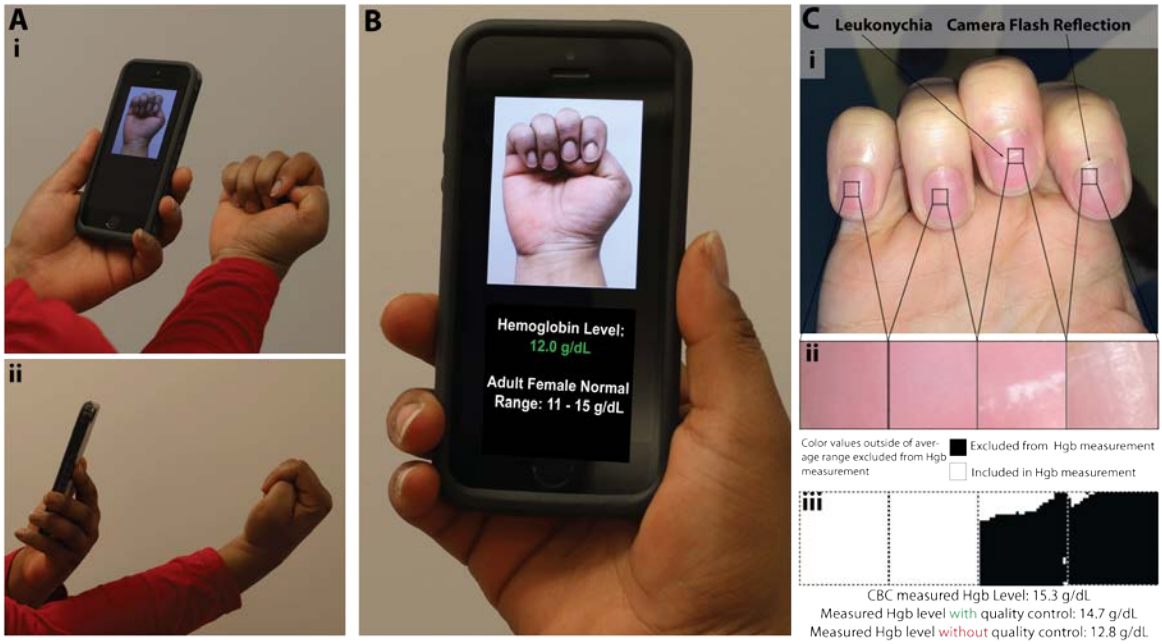

Figure 22: Example implementation of the smartphone image analysis system into a smartphone app enables non-invasive, patient-based measurement of blood Hgb levels via patient-sourced photos using only native hardware of the smartphone itself. A-C) proposed implementation of this system integrated into a smartphone app. **A i-ii)** As a patient obtains a smartphone photo of his/her fingernail beds, and without the need for any blood sampling or additional smartphone attachments or external calibration tools, **B)** the smartphone image analysis app quantitatively measures blood Hgb levels. **C i)** As smartphone images with fingernail irregularities such as camera flash reflections or leukonychia may affect Hgb level measurements, quality control measures of the Hgb measurement app detects and omits those irregularities to preserve measurement integrity and accuracy. **C ii)** To that end, the app randomly selects regions of interest from within the fingernail image, **C iii)** and any color values that fall outside of expected color ranges are excluded from Hgb measurement. In this example, when the quality control system was implemented to exclude the fingernail bed irregularities, Hgb level was measured to be 14.7 g/dL, comparable to the patient's CBC Hgb level of 15.3 g/dL. When the quality control was not implemented, Hgb level was measured at 12.8 g/dL. The quality control algorithm therefore resulted in a 76% reduction in error. Note that as the smartphone image-based algorithm is device-agnostic, the analysis of the smartphone images could also be transmitted to another device (e.g. laptop, cloud-based server) for remote analysis.

5.2.2 Skin Tone and Lighting Condition Analysis

Automated analysis of skin tone was employed to measure the effect that a subject's skin tone had in the algorithm's ability to accurately measure Hgb levels. The

CIELAB color space was designed as a device-independent color space capable of approximating human vision[94]. Much like the human brain, this color space distinguishes between light and dark ("L" value), between red and green ("A" value), and between blue and yellow ("B" value). A patch of skin on the middle finger was selected from each image, and the RGB color values in this region were transformed into the CIELAB colorspace. The lightness, or "L*", value most accurately represents human perception of skin color and serves as a n ideal variable linearly modeling skin tones in a population on a scale of dark skin to light skin[95]. The average "L*" value was found within the skin patch via a custom MATLAB script and compared with the residual for each image to establish the effect that skin tone (dark vs light) has on Hgb measurement accuracy (Fig. S2B). Furthermore, the impact of background lighting conditions was analyzed by recording background brightness via a digital light meter. Impact of background brightness on Hgb measurement was determined by comparison of background brightness with residuals.

5.2.3 Phone parameter analysis

A single subject's fingernail beds were imaged using multiple makes and models of smartphones (Apple, Cupertino, CA; Motorola, Schaumburg, IL). Each image was taken in the same location in rapid succession (~15 sec between images) in order to minimize fluctuations due to Hgb variation over time and background lighting condition. Color pixel intensity data was recorded for each smartphone model.

5.2.4 Quality Control

5.2.4.1 Fingernail irregularity exclusion

Preliminary experiments were completed to develop image analysis algorithms capable of performing quality control on the images, allowing me to minimize the impact of fingernail or image irregularities. A sample image of an individual with leukonychia (also known as white nails) was taken, producing a camera flash reflection on some nails. ROIs were chosen to contain these imaging and fingernail irregularities, and software was written in MATLAB to exclude them. Color values in the ROI were excluded if they fell outside of the expected color range of fingernail beds. The expected color range was calculated as the standard deviation of the average pixel intensity color values for the study population.

5.2.4.2 Detecting Presence of fingernails

In addition to excluding irregularities, I was able to determine the presence of fingernails using the Vision application programming interface (API) (Google, Mountain View, CA). A script was written in Python using the IDLE IDE that sends images to the Google platform, makes requests of the cloud service, and receives and interprets results from the server. Initially, images of fingernail beds are sent to the server with a request. The request asks the server via the Vision API to characterize or label the images that it receives. Machine learning algorithms are used in conjunction with Google's vast repository of images to label the image based on the content of that image. The custom Python script receives, interprets, and then characterizes the result. The response from the server is typically a series of 5 labels, along with corresponding data regarding the degree

of confidence in that label being correct. This quality control looks for the label "finger" or "fingernail" and records its confidence, allowing me to determine whether or not a user has indeed taken an image of a fingernail bed. Incorrect labels and responses are also categorized.

5.2.5 Blood flow interference

Table 4: Hgb level measurement experiences little variation when subjects are subjected to increased heart rate or changes in hand temperature. When heart rate increased due to moderate and heavy exercise, Hgb level measurement increased by 0.19 g/dL and 0.78 g/dL relative to resting heart rate (75BPM) Hgb level measurement, respectively. When hand temperatures corresponding to exposure to ambient temperatures of 4^{o}C and 39^{o}C, Hgb level measurement increased by 0.62 g/dL and 0.52g/dL, respectively, relative to room temperature (24^{o}C) Hgb level measurement. n = 4 per case (resting heart rate, room temperature, etc.). In each case, this change is clinically insignificant. Throughout the course of this entire study, regardless of the conditions, the mean |error| = 1.08 g/dL, n = 24 measurements, which is nearly the same as the mean |error| of the large scale clinical study (± 1.04 g/dL).

	Mean Heart Rate			**Ambient Temperature**	
	100 BPM	**149 BPM**		**4^{o}C**	**39^{o}C**
Mean Hgb Level Change From Baseline (smartphone)	0.19 g/dL	0.78 g/dL		0.62 g/dL	0.52 g/dL

The impacts of increases in heart rate as well as fluctuations in hand temperature were examined, as both conditions alter blood flow and fluctuations in blood flow to the fingernail beds may potentially affect Hgb measurement[105,106]. Even when heart rate increased significantly due to moderate or heavy exercise (100 BPM and 149 BPM, respectively) Hgb measurements only slightly increased by 0.19 g/dL and 0.78 g/dL when compared to Hgb measurement at resting heart rate (75BPM). When hand

temperature fluctuated between -4C or +3C of room temperature levels, corresponding to exposure to ambient temperatures of 4^{o}C and 39^{o}C, Hgb measurements increased by 0.62 g/dL and 0.52 g/dL, respectively, relative to Hgb measurement at room temperature (24^{o}C). However, in either case, the increase was not clinically significant (Table 4).

5.2.6 Heart Rate interference

Additional subjects' (n = 4) resting heart rates were measured and their fingernail beds were imaged at resting heart rate. Subjects were then instructed to conduct mild exercise (walking) and, subsequently, heavy exercise (jumping jacks) for one minute in order to increase and record heart rate. Fingernail beds were imaged after mild and heavy exercise. A blood draw was conducted and Hgb levels were determined with a CBC. Images from each activity state were used to measure Hgb via the image analysis algorithm and to measure the impact that heart rate has on smartphone Hgb measurement.

5.2.7 Skin Temperature Interference

Additional subjects' (n = 4) hands were dipped in cooled or heated water until a temperature change of -4^{o}C or +3^{o}C was recorded, respectively. Fingernail beds were imaged after each condition, as well as prior to temperature change, to establish a baseline. A blood draw was conducted, and Hgb levels were determined with a CBC. Images from each hand temperature condition were used to measure Hgb via the image analysis algorithm and determine the impact that skin temperature has on smartphone Hgb measurements.

5.2.8 Repeatability

Ten images were taken of an individual's fingernails, and Hgb was measured via the image analysis algorithm on each image. Repeatability was reported as the standard deviation of the measured Hgb levels.

5.2.9 Camera color calibration

5.2.9.1 Color calibration scheme

Camera calibration was used in order to facilitate agreement between the Hgb readouts between multiple models of smartphones, as different phones use different cameras. Images of the same subject matter (Figure 23) were taken under identical conditions and on multiple different models of smartphones including the iPhone 5S, iPhone 7 (Apple, Cupertino CA), and Samsung Galaxy S7 (Samsung, Seoul, South Korea). The subject matter in these images consisted of a collage of paint swatches (The Home Depot, Atlanta GA) arranged in a grid containing different colors spanning the entire RGB color space. RGB data was extracted from regions within each of these grids on each image from the multiple smartphone models. Since the Hgb measurement algorithm was developed using images taken from an iPhone 5S, color data from images taken with the other smartphone models was compared to the color data taken from the iPhone 5S in order to develop a color transformation that relates RGB information of the desired smartphone to RGB information of the original smartphone.

5.2.9.2 RGB transform

The RGB transform function was developed by first collecting RGB data from each colored region in the color calibration card within images taken on each camera type. For each color (R, G, B) linear regression was used with the RGB values from iPhone 5S (the original phone the algorithm was developed on) as the independent variables and the RGB values from the other phones as the dependent variables (Figure 24). The equation of the resulting regression lines were used to transform RGB values of

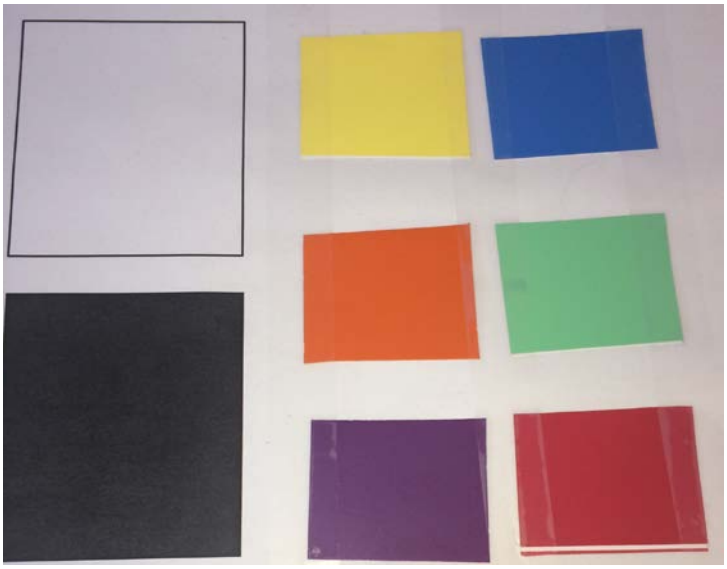

Figure 23: Color calibration card design. Images of the above color calibration card were taken using multiple smartphone models, and the differences in RGB color data from each color region within the image were used to develop an RGB calibration transformation for each smartphone model.

the images taken on new phones so that they matched the RGB values of the original phone. Linear regression and r^2 values were calculated with Microsoft Excel (Microsoft, Redmond, WA).

5.2.10 Physician vs app performance comparison

Testing was conducted to compare the diagnostic power of physicians against the diagnostic power of the hemoglobin measurement app. A panel of 5 trained and board-certified hematologists from CHOA, Emory University, and Grady Hospital measured

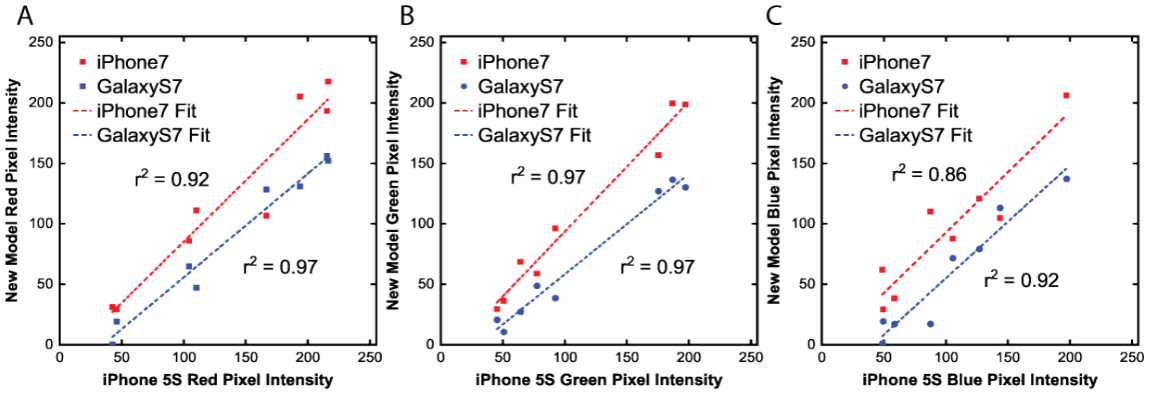

Figure 24: Linear regression was used to correlate RGB values extracted from images taken by multiple smartphones to RGB values taken from iPhone 5S images. Red **(A)**, Green **(B)**, and Blue **(C)** pixel intensity values were extracted from each box in the color calibration card. Each data point represents the RGB value taken from a single square on a given phone. Values derived from the iPhone 5S (the original phone the Hgb measurement algorithm was developed on) are plotted on the X-axis and values derived from new, prospective phones are plotted on the Y-axis. Linear regressions were calculated for each prospective phone against the iPhone 5S and the resulting equation displayed relates RGB values from each new phone (iPhone 7 – Top, Galaxy S7 - Bottom) to RGB values of the original phone.

Hgb levels and their diagnostic accuracy was compared against that of the app. Accuracy testing was conducted in two phases: 1) measuring Hgb using images of subjects' fingernails, and 2) using Hgb measurements from in-person physical examinations of subjects' clinical pallor. Diagnostic accuracy for comparison was determined by CBC testing. Each subject had a blood venous draw and a CBC was conducted using a clinical hematology analyzer (Advia 2120i, Siemens, Berlin, Germany).

5.2.10.1 Hemoglobin measurement via physical examination

These physicians were told to inspect the fingernails of 22 subjects and measure their hemoglobin levels. A research associate then used the app to measure these subjects Hgb levels. These subjects' ages ranged from 25 to 62 years old. In order to minimize the possibility of other physical factors subconsciously influencing the analysis of the physicians, a second study was conducted using image of subjects' fingernails.

5.2.10.2 Hemoglobin measurement from images of fingernails

Images were taken of 50 subjects fingernails from the previously described clinical study. These subjects' ages ranged from 1 to 62 years old. Physicians were instructed to analyze each image and measure Hgb levels. For comparison, images were loaded into the app, and the Hgb measurement protocol was performed on these images. All images and analysis were taken using an iPhone 5S. It is important to note that these images were not used in the development of the underlying image analysis algorithm.

5.2.11 Study Approval

The study regarding the relation of finger temperature and exercise to measurement error was approved by the Georgia Institute of Technology IRB (approval number H17118)

5.3 Results and Discussion

5.3.1 Quality control measures eliminate error caused by fingernail irregularities and ensure presence of fingernails in each image taken.

I have implemented quality control measures into the app to enable the app to perform better when presented with challenges such as irregular fingernail beds and user error. Quality control software has the ability to exclude common fingernail irregularities from Hgb calculation. This is done by excluding fingernail regions that are significantly lighter or significantly darker than the average fingernail bed. These abnormally light or dark regions correspond to nailbed irregularities such as bruising due to injury, leukonychia, melanonychia, camera flash and external light reflection, or side effects of certain medications (e.g. chemotherapy drug treatments can lead to irregular striping of the nailbed[107]). Preliminary results are promising, indicating that Hgb level measurement improves to 14.7 g/dL up from 12.8 g/dL (CBC Hgb = 15.3 g/dL) in a single patient when quality control software is used (Figure 22C). I hypothesize that the subject's leukonychia and camera flash reflection off the surface of the fingernail are incorrectly being interpreted as clinical pallor by the algorithm, leading to artificially low Hgb level measurement. The impact of these irregularities seems to be mitigated by quality control software, indicating that quality control is essential to ensuring that true Hgb level measurements are obtained. Furthermore, quality control characterization utilizing Google's cloud computing service and Vision API is able to effectively and efficiently characterize fingernail images by quickly determining the presence of fingernails in an image. For this technology to achieve its full potential, it will need to be implemented into a smartphone app that patients can download and use at home, and this home-based

use model will likely introduce some user error that was not seen when testing was conducted by an expert user of this technology. I pre-emptively addressed the most basic usability concern by determining whether a fingernail was present at all within an image. Each image taken during the clinical study was process by the Google Vision API via the Google cloud service. 332 of 337 images analyzed were characterized as containing a "finger" or "fingernail" (Figure 25). Furthermore, approximately 97% of images analyzed were characterized as containing these objects with a degree of certainty of greater than 85%. If implemented into the app, this result will enable the app to ensure that a user has indeed captured an image of their fingernails sufficient to conduct Hgb measurement, eliminating a major potential source of user error. This should significantly improve app performance when used at home.

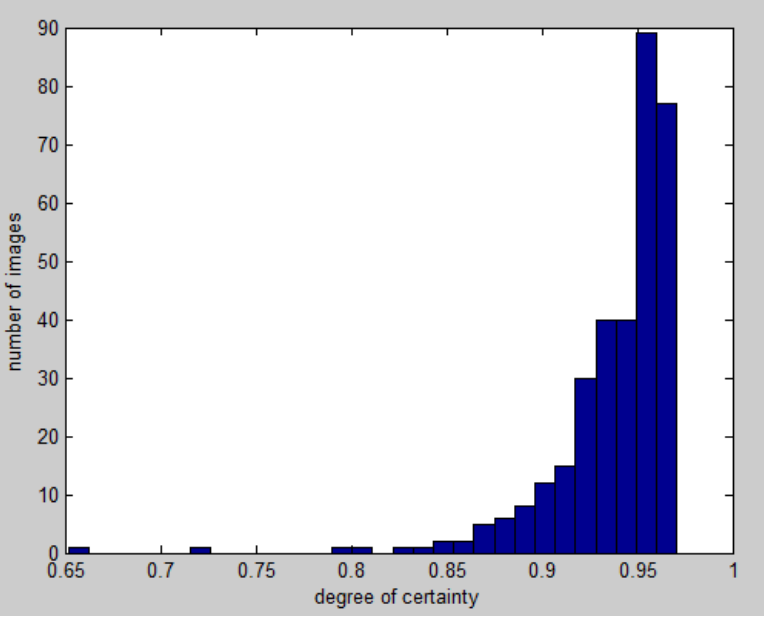

Figure 25: Degree of certainty that a fingernail is present in clinical study photos. Each response from Google Vision API containing the word "finger" or "fingernail" was recorded. All but 5 clinical study images returned one of these words. ~97% of responses were characterized as containing a finger or fingernail with a degree of certainty of > 0.85. N = 337 images

5.3.2 *Images taken of subjects' fingernail beds are consistent amongst smartphone images taken with different makes and models of smartphones when color calibration strategies are utilized.*

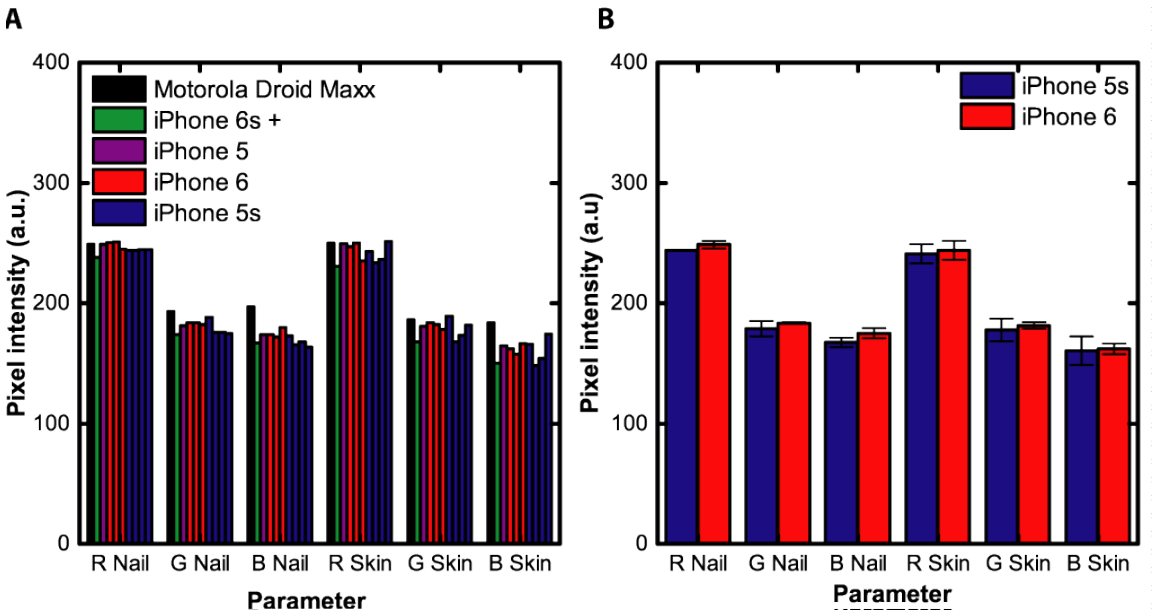

Figure 26: The image analysis algorithm is camera-agnostic. A) RGB values of both fingernails and the skin of within a sample image taken with different models of smartphones remain consistent. n = 10 phones. **B)** There was no statistically significant difference between two different smartphone models developed by the same manufacturer ($P > 0.05$). Statistical analysis was performed with a student's t-test assuming unequal variance. n = 3 phones/model. Error bar reported as standard deviation

This algorithm must be easily accessible by the smartphones used to image subjects' fingernails in order to maximize the impact that this image analysis anemia prediction algorithm can make. This can best be accomplished in the form of an on-board smartphone application. The smartphone market is large and diverse. An application relying on data extracted from images taken by a smartphone must be camera-agnostic to ensure that the same readings are given regardless of the make or model of the camera used to take pictures. Preliminary studies have been conducted to ensure cameras from different mobile phones produce similar images of subjects' fingernails. Images of a

single subject's fingernail beds were taken under identical imaging conditions using multiple smartphone models and the results were compared. Images were taken using phones made from different manufactures, including Apple and Motorola, as well as different models of phones made by the same manufacturer (i.e. Apple iPhone 4, 5, 5s, 6, 6s+). I have shown that RGB pixel intensity values of the subject's fingernail beds and a control patch of skin are remarkably similar between images taken with different smartphones (Figure 26A). Additionally, I show no statistically significant difference ($p > 0.05$) between pixel intensity values of images of fingernail beds taken by two different models of Apple iPhones (Figure 26B), with the exception of the r channel pixel intensity value of the fingernail bed ($p = 0.03$). However, this difference is very small (~4 A.U.), contributing to a difference in Hgb level estimation of approximately 0.2g/dL).

However, further studies show that this pixel intensity agreement not the case

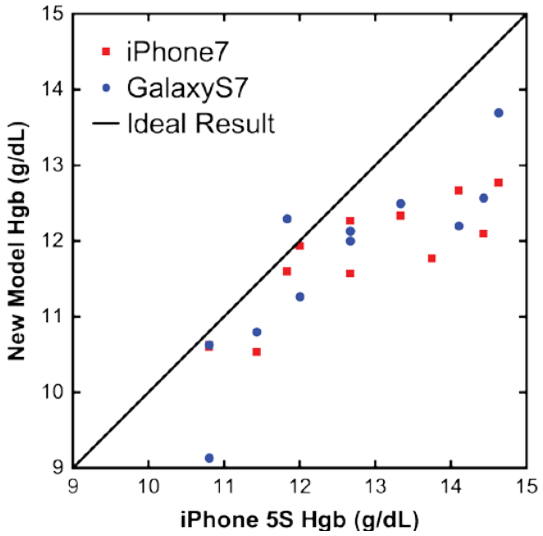

Figure 27: Color calibration facilitates agreement between alternate smartphone models. When the color calibration protocol is used, alternate models of smartphones (iPhone 7 and Samsung Galaxy S7) display similar results. The app was tested on the same group of subjects in each case. An offset from the ideal result is present (N = 11).

with newer models of Samsung and Apple phones. Thus, color calibration was necessary to facilitate pixel intensity agreement between images taken on the different models of smartphones. Taking images of color arrays with the different smartphone models allowed me to develop a pixel intensity value transformation to relate the data from each smartphone to ensure Hgb measurement accuracy (Figure 23, Figure 24). Implementation of this transformation facilitates agreement between the alternate smartphone models (iPhone 7 and Samsung Galaxy S7), however, some experimental bias has been introduced that causes a slight offset from the iPhone 5S Hgb measurement baseline that must be corrected (Figure 27).

5.3.3 Smartphone Image-based Algorithm Performance to Potential Sources of Interference and Variability

Use of fingernail beds as the imaging area is ideal due to the fact that fingernail beds contain minimal amounts of melanin compared to other parts of the skin[67,68], theoretically enabling this technique to be insensitive to subject skin tone. Preliminary results presented earlier seem to support this hypothesis (Figure 12). To further address this experimentally using the most accurate Hgb measurement algorithm, images were converted into the CIELAB color space, a commonly used color quantification system that quantifies color as perceived by the human eye[67]. In particular, the L* value in this color space has been shown to serve as a linear indicator of skin tone[67]. The relationship between the subjects' skin tones and Hgb measurement residuals was determined by measuring the L* value of a patch of skin adjacent to the fingernail. L* did not correlate ($r = 0.17$) with Hgb measurement residual, indicating that subject skin tone has little impact on the ability of this system to measure Hgb levels (Figure 28A).

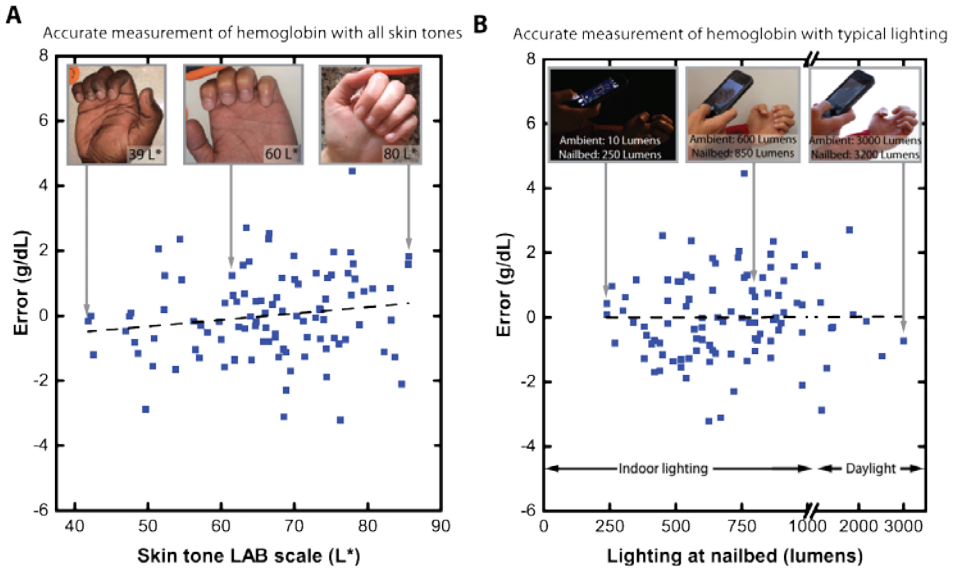

Figure 28: Background lighting and subject skin tone have minimal effect on Hgb measurement accuracy. Plotting measurement error against **(A)** skin tone and **(B)** background lighting reveals low and negligible correlation (r = 0.17 and r = 0.004 respectively) in either case. Dashed lines indicate linear fit between the measurement error and the tested parameter (skin tone and background lighting respectively). Inset images highlight a representative range of measured background skin tones and lighting conditions. n = 100 patients.

For accessibility in dynamic clinical settings, the smartphone Hgb level estimation technology must function under a wide variety of background lighting conditions. To that end, using luminous flux readings on a digital light meter, no correlation (r = 0.004) was found between room brightness and Hgb measurement residual, indicating that this method can be used in a wide variety of settings and lighting conditions (Figure 28B). Use of the camera flash resulted in the most accurate Hgb level measurement, likely due to the normalization of background lighting conditions provided by the camera flash (Figure 29). Furthermore, ensuring that the technology is agnostic to the smartphone make and model, RGB pixel intensity values of the subject's fingernail beds and a control patch of skin were found to be similar between images taken with smartphones made from different manufacturers and models (Figure 26A). Additionally,

no statistically significant difference existed between pixel intensity values of fingernail bed images obtained by two different smartphone (Figure 26B). Finally, the precision of Hgb level measurements using this technology was found to be ±0.17 g/dL when tested on multiple images of the same individual's fingernails.

5.3.4 Color calibration facilitates agreement in Hgb measurement between different smartphone models

While preliminary data showed that there is little variability between smartphone images, the limited variability is sufficient to cause disagreement in Hgb measurement between different smartphone models when the image analysis algorithm is integrated into smartphone apps. In order to resolve this disagreement, I developed a color calibration protocol. In this protocol, images are taken of multiple standardized colors and the RGB color data is compared between the different phones. Using the iPhone 5S as the standard and independent variable in linear regression analysis, an equation is developed that can be used to transform RGB values of other smartphone models to fit the Hgb measurement model designed using iPhone 5S images (Figure 27). This protocol can be adapted to conduct color calibration on any future smartphone models that become available and in which compatibility with this app is desired.

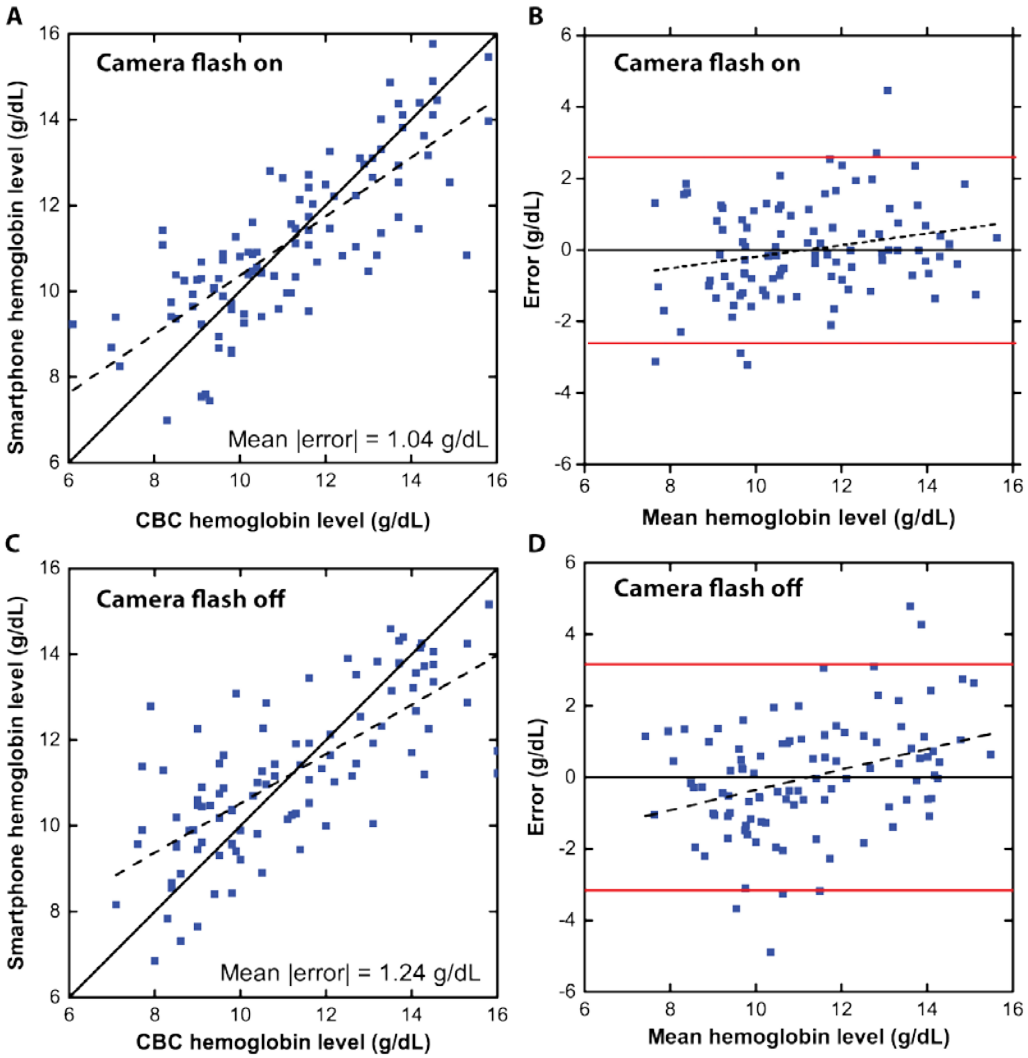

Figure 29: Camera flash improves performance of the Hgb measurement algorithm. Use of the camera flash **(A-B)** leads to a decrease in the Hgb measurement error, as well as an increase in the correlation between the smartphone and CBC measurements, compared to when no camera flash is used (Flash on: r = 0.23 Flash off: r = 0.35) **(C-D)**. Bland-Altman analysis shows a slight decrease in the correlation between the residual and the average Hgb level between the two tests when the camera flash is used. This indicates the presence of some experimental bias that is mitigated by use of the camera flash. n = 100

5.3.5 The app outperforms trained hematologists's ability to measure Hgb levels via physical examination.

Hematologists were challenged to measure Hgb levels in patients via physical examination of the fingernails as well as inspection of images of fingernails. When

physical exams were conducted, there was no difference between the diagnostic performance of trained haematologists (1.76 g/dL average error) vs the app (1.83 g/dL average error) in this preliminary study. I hypothesize that this could have been the case due to the Hgb distribution of the study subject pool. In the subject pool for physical examinations (n = 22), Hgb levels ranged from 10.7 g/dL to 16.9 g/dL. Thus, none of the

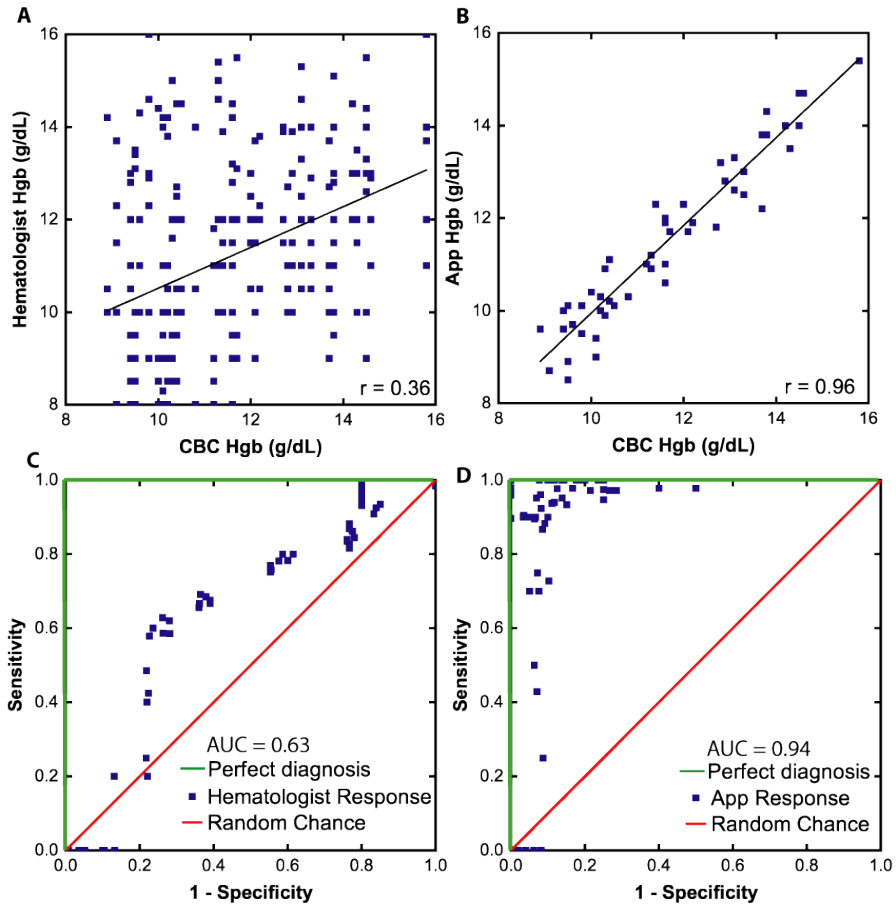

Figure 30: Smartphone app outperforms hematologist estimations when diagnosing Hgb using images of fingernails. A) Pooled results of 5 trained hematologist's Hgb measurements after inspection of fingernail images. Measurements were accurate to within ± 4.5 g/dL (95% CI, r^2 = 0.13, n = 50 subjects, N = 5 hematologists). **B)** Results of Hgb measurement via the app. Measurements were accurate to within ± 1.0 g/dL (95% CI, r^2 = 0.91, n = 50 subjects). **C, D)** ROC curves of the haematologist response (AUC = 0.63) are much closer than the app responses (AUC = 0.94) to the random chance line, indicating that the App has a greater diagnostic accuracy than trained haematologists.

subjects were anemic, leading to the app and physicians to measure similar Hgb levels. Furthermore, it is likely that the physicians displayed some bias when measuring Hgb, as they had no reason to expect that these subjects were anemic (the subjects were identified and consented from a research lab at the Georgia Institute of Technology rather than a clinical setting. This could have potentially led to artificially high Hgb results.

In order to account for physician bias and include a wider range of Hgb levels to more accurately reflect the general population who could potentially use this app, images of patients were used for comparison of the diagnostic capabilities of physicians and the app. When looking at images of fingernail beds, the hematologists measured Hgb levels to within ± 4.5 g/dL (95% CI, $r^2 = 0.13$, n = 50, Figure 30A). When the app was used to measure Hgb on the same images, the app was able to measure Hgb to within ± 1 g/dL (95% CI, $r^2 = 0.91$, n = 50, Figure 30B). This represents a significant improvement and proves that the image analysis algorithm translates well upon integration into a smartphone app. In fact, there is an increase in diagnostic accuracy when the app is used, likely due to the quality control software that has been implemented.

5.4 Conclusions and Future Approaches

It is possible that this Hgb estimation technique will not work under extreme lighting conditions. It is necessary to conduct experiments to determine the lighting condition boundaries under which the algorithm successfully determines Hgb levels. This will be done by measuring the error in Hgb level estimation of the same subject in various lighting conditions exceeding the conditions typically found in clinical settings. These settings represent potential field use of this technology. This is an important

experiment to conduct, as the primary value proposition of this technology to the user is based on the ability to conduct Hgb testing away from clinical settings.

In order to develop a truly POC diagnostic tool, the Hgb estimation algorithm must be incorporated into a mobile platform and used by individuals with, or at risk for, anemia. In order to ensure that this diverse user base can accurately measure their Hgb levels, a "quality control" system must be in place to ensure that the images being taken by these users are of sufficient quality to generate accurate Hgb level results. While preliminary quality control systems have already been developed, further algorithms will be develop which account for other common irregularities in images of fingernails will be developed. In addition to the abnormalities studied, image irregularities that could lead to inaccurate Hgb level estimation include presence of abnormal fingernail bed pigmentation, abnormal imaging brightness, and lack of image focus. These algorithms will serve as an extension of the preliminary quality control algorithms already developed. Convolution of fingernail bed images with edge detection kernels[108] will be used to detect edges within the fingernail beds corresponding to abrupt color changes caused by abnormal fingernail bed pigmentation for each finger. Data from fingernail beds with irregular pigments and irregularities will be removed from consideration during Hgb level measurement. An acceptable range of background brightness will be established using images taken as part of the clinical study which have been shown to allow for accurate Hgb estimation. The quality control algorithm will simply compare the background brightness to the reference standard to determine whether or not to reject the image. The quality control algorithm will reject out-of-focus images as well. Determination of the "sharpness" of image focus can be determined in conjunction with

the edge detection function previously described. The intensity of edges will determine how sharp, or in focus, they are. A sharpness of edges metric will be determined using images that have been shown to yield an accurate Hgb level estimation. This metric will be used to determine rejection criteria for image focus. This “quality control” algorithm will be used to reject images that could potentially yield inaccurate Hgb estimation results and will prompt the user to take higher quality images.

As I have presented in this chapter, there are significant differences in how images are interpreted across different models and manufacturers of smartphone. The color correction protocol must serve as a blueprint for enabling the app to function properly on smartphones that have not currently been tested. Further testing must be conducted to optimize this protocol, as there exists some disagreement in Hgb levels measured on more advanced smartphone models compared to the model used to originally develop the Hgb measurement algorithm. Furthermore it is possible that requiring subjects to manually select regions of interest within images of their fingernail beds will prove difficult, causing inaccurate Hgb estimation. If this is the case, it may be necessary to build in an image analysis feature that extracts regions of interest automatically, which is an option that I have explored via utilization of convolutional neural networks that have been used in computer vision. Overall, I am able to report successful preliminary implementation of this image analysis algorithm into multiple mobile platforms across different smartphone manufacturers. As this technology in its current form is capable of outperforming trained hematologists in anemia diagnosis from physical examinations of fingernail bed images, this app represents a major step in the

development of a truly translational, noninvasive, convenient, and inexpensive diagnostic tool for anemia.

CHAPTER 6. CONCLUDING REMARKS AND FUTURE DIRECTIONS

Given the performance of this technology in the general population, this completely noninvasive technology that requires only photos obtained from smartphones has significant implications as a screening tool in the general population, as anemia is prevalent worldwide, afflicting nearly two billion people, especially young children, the elderly, and pregnant women. The ability to inexpensively diagnose anemia with high sensitivity, completely noninvasively, and without the need for any external smartphone attachments or calibration equipment represents a significant improvement over current POC anemia screening. Optimizing sensitivity is of paramount importance for a screening tool, due to the ability to correctly identify a high percentage of anemia cases even if this negatively impacts specificity. In its current form, this technology requires the user to simply obtain a fingernail image, which can then be analyzed with an on-board smartphone app that comprises a novel image analysis algorithm to output the Hgb level measurement or be transmitted remotely to another device (e.g. laptop, desktop computer, or cloud-based server with my algorithm embedded into their systems) for remote analysis, the results of which can be immediately transmitted back to the user. After identifying subjects that may possibly be anemic, either type of system can recommend confirmatory Hgb level testing with a CBC, allowing any false positives to avoid unnecessary treatment. Given the ever-increasing rate and near ubiquity of smartphone ownership worldwide[9], this noninvasive, inexpensive, patient-operated Hgb measurement algorithm allows those at risk of anemia to monitor their conditions using

only the native hardware included with their own smartphone. This is particularly pertinent in low resource settings where, contradictory to the relative lack of medical infrastructure, mobile phone networks are quite extensive and have "leapfrogged" landlines[109].

Additionally, this system has the potential to fundamentally alter the management of patients with chronic anemia. During the course of several weeks, a patient may take images of their fingernail beds and enter their CBC-measured Hgb levels that were obtained as part of their regular outpatient clinical care. Results suggest that these images and Hgb levels may be used to "teach" the smartphone phone to develop a "calibration" personalized and tailored to each individual patient. In times of clinical stress, these patients, such as those with genetic causes of anemia or with cancer and are undergoing chemotherapy, may experience rapid, life-threatening, precipitous drops in Hgb and require constant monitoring to determine their need for transfusions. Using this technology, patients could potentially self-monitor their anemia from the comfort of their own home, rather than through inconvenient and recurring clinic visits. In addition, some patients with chronic anemia due to a genetic etiology require chronic transfusions to survive. These scheduled transfusions are currently administered at convenient and regular intervals, and not based on clinical need[110]. Hence, a patient may be transfused too early, exposing them to unnecessary transfusion-related effects (i.e. iron overload, risk of infection), while patients transfused too late may require urgent hospitalization if they develop symptomatic anemia or their Hgb levels decrease to a dangerous level. By enabling continuous and simple monitoring, this technique may empower patients and lead to better allocation of blood bank resources. Moreover, further data collection will

increase the size of the patient image pool, facilitating the incorporation of deep machine learning "Big Data" techniques to further refine the Hgb measurement algorithm[111].

Overall, this CBC-validated, smartphone image-based Hgb measurement technology has the potential to dramatically improve upon the accuracy, cost, and convenience of current Hgb measurement devices while also eliminating the need for anything other than a smartphone, representing a significant improvement over other POC Hgb measurement technologies[5,6,24,25,57,112,113]. With this smartphone image-based Hgb measurement system, any person – healthy or ill - in any location, at any time, now has access to an important health metric and may seek care accordingly. Moreover, healthcare officials in low resource settings may use this technology to efficiently and effectively allocate limited healthcare resources (e.g. transfusions, high-risk obstetrical services) and medications (e.g. nutritional supplementation such as iron, folate, or vitamin B12)[114] for the patients with the most severe anemia. This completely noninvasive, algorithm-based approach represents a paradigm shift in the way anemia can be screened, diagnosed, and monitored globally. As the system requires no reagents or equipment, the healthcare cost savings could also be significant.

Going forward, the potential for user error as well as inter/intra-smartphone variability leading to Hgb measurement error will be addressed in the form of a full clinical assessment, and I will also investigate the efficacy of the smartphone image-based algorithm in which patients will use this app as a self-test using multiple models and manufacturers of smartphones. This study will also evaluate and improve upon the quality control measures that I have implemented thus far. Overall, this system represents the first noninvasive system that can potentially replace blood-based testing without

requiring any additional medical equipment. While the current algorithm and app development focused specifically on anemia, this approach can potentially be adapted to quantitatively diagnose other conditions that manifest in physical exam abnormalities (i.e. jaundice and cyanosis)[29,30]. This technology has the capability to completely shift the treatment paradigm for anemia and other disorders that manifest in physical changes.

REFERENCES

1 World Health, O. Worldwide prevalence of anaemia 1993-2005: WHO global database on anaemia. *Worldwide prevalence of anaemia 1993-2005: WHO global database on anaemia.* (2008).

2 Galanello, R. & Origa, R. Beta-thalassemia. *Orphanet Journal of Rare Diseases* **5**, 1-15, doi:10.1186/1750-1172-5-11 (2010).

3 Frenette, P. S. & Atweh, G. F. Sickle cell disease: old discoveries, new concepts, and future promise. *The Journal of clinical investigation* (2007).

4 Karnad, A. & Poskitt, T. R. The Automated Complete Blood Cell Count: Use of the Red Blood Cell Volume Distribution Width and Mean Platelet Volume in Evaluating Anemia and Thrombocytopenia. *Archives of Internal Medicine* **145**, 1270-1272, doi:10.1001/archinte.1985.00360070150025 (1985).

5 Schenck, V. H. & Falkensson, M. Evaluation of" HemoCue," a new device for determining hemoglobin. *Clinical Chemistry* (1986).

6 Tyburski, E. A. *et al.* Disposable platform provides visual and color-based point-of-care anemia self-testing. *Journal of Clinical Investigation* **124**, 4387-4394, doi:10.1172/JCI76666 (2014).

7 Neufeld, L. *et al.* Hemoglobin measured by Hemocue and a reference method in venous and capillary blood: a validation study. *Salud Pública de México* **44**, 219-227, doi:10.1590/S0036-36342002000300005 (2002).

8 Gubala, V., Harris, L. F., Ricco, A. J., Tan, M. X. & Williams, D. E. Point of care diagnostics: status and future. *Anal Chem* **84**, 487-515, doi:10.1021/ac2030199 (2012).

9 GSMA. The Mobile Economy. *The Mobile Economy* (2016).

10 Wolf, J. A. *et al.* Diagnostic Inaccuracy of Smartphone Applications for Melanoma Detection. *JAMA Dermatology* **149**, 422-426, doi:10.1001/jamadermatol.2013.2382 (2013).

11 Rappaport, K. M. *et al.* Assessment of a Smartphone Otoscope Device for the Diagnosis and Management of Otitis Media. *Clin Pediatr (Phila)* **55**, 800-810, doi:10.1177/0009922815593909 (2016).

12 Breslauer, D. N., Maamari, R. N., Switz, N. A., Lam, W. A. & Fletcher, D. A. Mobile Phone Based Clinical Microscopy for Global Health Applications. *PLoS ONE* **4**, doi:10.1371/journal.pone.0006320 (2009).

13 Chalco, J. P., Huicho, L., Alamo, C., Carreazo, N. Y. & Bada, C. A. Accuracy of clinical pallor in the diagnosis of anaemia in children: a meta-analysis. *BMC Pediatr* **5**, 46, doi:10.1186/1471-2431-5-46 (2005).

14 Kalantri, A., Karambelkar, M., Joshi, R., Kalantri, S. & Jajoo, U. Accuracy and Reliability of Pallor for Detecting Anaemia: A Hospital-Based Diagnostic Accuracy Study. *PLoS ONE* **5**, doi:10.1371/journal.pone.0008545 (2010).

15 Sheth, T. N., Choudhry, N. K., Bowes, M. & Detsky, A. S. The relation of conjunctival pallor to the presence of anemia. *J Gen Intern Med* **12**, 102-106, doi:10.1046/j.1525-1497.1997.00014.x (1997).

16 Thaver, I. H. & Baig, L. Anaemia in children: Part I. Can simple observations by primary care provider help in diagnosis? *JPMA. The Journal of the Pakistan Medical Association* **44**, 282-284 (1994).

17 Weber, M. W. *et al.* Pallor as a clinical sign of severe anaemia in children: an investigation in the Gambia. *Bulletin of the World Health Organization* **75 Suppl 1**, 113-118 (1997).

18 Deinard, A. S., List, A., Lindgren, B., Hunt, J. V. & Chang, P. N. Cognitive deficits in iron-deficient and iron-deficient anemic children. *J Pediatr* **108**, 681-689, doi:10.1016/S0022-3476(86)81041-1 (1986).

19 Eschbach, J. W. The anemia of chronic renal failure: pathophysiology and the effects of recombinant erythropoietin. *Kidney international* (1989).

20 Groopman, J. E. & Itri, L. M. Chemotherapy-induced anemia in adults: incidence and treatment. *Journal of the National Cancer …*, doi:10.1093/jnci/91.19.1616 (1999).

21 Volberding, P. Consensus statement: Anemia in HIV infection—current trends, treatment options, and practice strategies. *Clinical Therapeutics* **22**, 1004-1020, doi:10.1016/S0149-2918(00)80081-8 (2000).

22 World Health, O. Micronutrient Deficiencies. *Micronutrient Deficiencies*.

23 McMurtry, M. C., Noel, M., Chambers, C. T. & McGrath, P. J. Children's fear during procedural pain: Preliminary investigation of the Children's Fear Scale. *Health Psychology* **30**, 780, doi:10.1037/a0024817 (2011).

24 Collings, S. *et al.* Non-Invasive Detection of Anaemia Using Digital Photographs of the Conjunctiva. *PLOS ONE* **11**, doi:10.1371/journal.pone.0153286 (2016).

25 Suner, S., Crawford, G., McMurdy, J. & Jay, G. Non-invasive determination of hemoglobin by digital photography of palpebral conjunctiva. *The Journal of emergency medicine* **33**, 105-111, doi:10.1016/j.jemermed.2007.02.011 (2007).

26 Dorsey, R. E. & Topol, E. J. State of Telehealth. *The New England Journal of Medicine* **375**, 154-161, doi:10.1056/NEJMra1601705 (2016).

27 Maamari, R. N., Ausayakhun, S., Margolis, T. P., Fletcher, D. A. & Keenan, J. D. Novel telemedicine device for diagnosis of corneal abrasions and ulcers in resource-poor settings. *JAMA ophthalmology* **132**, 894-895 (2014).

28 Whitaker B.A., S. K., Schulman J., Green J. . The 2009 National Blood Collection and Utilization Survey. *US Department of Health and Human Services* (2011).

29 McGrath, J. S., Datir, S. & O'Brien, F. Why so blue? A case of neonatal cyanosis due to congenital methaemoglobinaemia (HbM Iwate). *BMJ case reports* **2016**, doi:10.1136/bcr-2016-216805 (2016).

30 Hyperbilirubinemia, S. o. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. *Pediatrics* (2004).

31 Goodnough, L. T. & Schrier, S. L. Evaluation and management of anemia in the elderly. *Am J Hematol* **89**, 88-96, doi:10.1002/ajh.23598 (2014).

32 Stevens, G. A. *et al.* Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995-2011: a systematic analysis of population-representative data. *Lancet Glob Health* **1**, e16-25, doi:10.1016/S2214-109X(13)70001-9 (2013).

33 Weiss, G. & Goodnough, L. T. Anemia of chronic disease. *N Engl J Med* **352**, 1011-1023, doi:10.1056/NEJMra041809 (2005).

34 Camaschella, C. Iron-Deficiency Anemia. *N Engl J Med* **373**, 485-486, doi:10.1056/NEJMc1507104 (2015).

35 Camaschella, C. New insights into iron deficiency and iron deficiency anemia. *Blood Rev* **31**, 225-233, doi:10.1016/j.blre.2017.02.004 (2017).

36 Goodnough, L. T., Nemeth, E. & Ganz, T. Detection, evaluation, and management of iron-restricted erythropoiesis. *Blood* **116**, 4754-4761, doi:10.1182/blood-2010-05-286260 (2010).

37 Kassebaum, N. J. *et al.* A systematic analysis of global anemia burden from 1990 to 2010. *Blood* **123**, 615-624, doi:10.1182/blood-2013-06-508325 (2014).

38 van Santen, S. *et al.* Hematologic parameters predicting a response to oral iron therapy in chronic inflammation. *Haematologica* **99**, e171-173, doi:10.3324/haematol.2014.106799 (2014).

39 Onken, J. E. *et al.* A multicenter, randomized, active-controlled study to investigate the efficacy and safety of intravenous ferric carboxymaltose in patients with iron deficiency anemia. *Transfusion* **54**, 306-315, doi:10.1111/trf.12289 (2014).

40 Vadhan-Raj, S., Dahl, N. V., Bernard, K., Li, Z. & Strauss, W. E. Efficacy and safety of IV ferumoxytol for iron deficiency anemia in patients with cancer. *J Blood Med* **8**, 199-209, doi:10.2147/JBM.S138474 (2017).

41 Vadhan-Raj, S. *et al.* Efficacy and safety of IV ferumoxytol for adults with iron deficiency anemia previously unresponsive to or unable to tolerate oral iron. *Am J Hematol* **89**, 7-12, doi:10.1002/ajh.23582 (2014).

42 Dhaliwal, G., Cornett, P. A. & Tierney, L. M., Jr. Hemolytic anemia. *Am Fam Physician* **69**, 2599-2606 (2004).

43 Chadburn, A. The spleen: anatomy and anatomical function. *Semin Hematol* **37**, 13-21 (2000).

44 Kaul, D. K., Fabry, M. E. & Nagel, R. L. Microvascular sites and characteristics of sickle cell adhesion to vascular endothelium in shear flow conditions: pathophysiological implications. *Proceedings of the National Academy of Sciences* **86**, 3356-3360, doi:10.1073/pnas.86.9.3356 (1989).

45 Platt, O. S., Thorington, B. D. & Brambilla, D. J. Pain in sickle cell disease: rates and risk factors. … *England Journal of* … (1991).

46 Perrotta, P. L. & Snyder, E. L. Non-infectious complications of transfusion therapy. *Blood Rev* **15**, 69-83, doi:10.1054/blre.2001.0151 (2001).

47 Petz, L. D. Drug-induced autoimmune hemolytic anemia. *Transfus Med Rev* **7**, 242-254 (1993).

48 Schwartz, R. S. & Costea, N. Autoimmune hemolytic anemia: clinical correlations and biological implications. *Semin Hematol* **3**, 2-26 (1966).

49 Barrett-Connor, E. Anemia and infection. *Am J Med* **52**, 242-253 (1972).

50 Rossifanelli, A., Antonini, E. & Caputo, A. Hemoglobin and Myoglobin. *Adv Protein Chem* **19**, 73-222 (1964).

51 Triplett, D. A. Coagulation and bleeding disorders: review and update. *Clin Chem* **46**, 1260-1269 (2000).

52 Ranucci, M. *et al.* Major bleeding, transfusions, and anemia: the deadly triad of cardiac surgery. *Ann Thorac Surg* **96**, 478-485, doi:10.1016/j.athoracsur.2013.03.015 (2013).

53 Kouides, P. A. Females with von Willebrand disease: 72 years as the silent majority. *Haemophilia* **4**, 665-676 (1998).

54 Mannucci, P. M. & Tuddenbam, E. G. The hemophilias: progress and problems. *Semin Hematol* **36**, 104-117 (1999).

55 Rockey, D. C. Occult gastrointestinal bleeding. *Gastroenterol Clin North Am* **34**, 699-718, doi:10.1016/j.gtc.2005.08.010 (2005).

56 Jaggernath, M. *et al.* Diagnostic Accuracy of the HemoCue Hb 301, STAT-Site MHgb and URIT-12 Point-of-Care Hemoglobin Meters in a Central Laboratory and a Community Based Clinic in Durban, South Africa. *PLoS One* **11**, e0152184, doi:10.1371/journal.pone.0152184 (2016).

57 Wang, E. *et al.* HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. *HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras*, 593-604, doi:10.1145/2971648.2971653 (2016).

58 Moore, L. J. *et al.* Evaluation of noninvasive hemoglobin measurements in trauma patients. *Am J Surg* **206**, 1041-1047, doi:10.1016/j.amjsurg.2013.08.012 (2013).

59 Paddle, J. J. Evaluation of the Haemoglobin Colour Scale and comparison with the HemoCue haemoglobin assay. *Bull World Health Organ* **80**, 813-816 (2002).

60 Marn, H. & Critchley, J. A. Accuracy of the WHO Haemoglobin Colour Scale for the diagnosis of anaemia in primary health care settings in low-income countries: a systematic review and meta-analysis. *Lancet Glob Health* **4**, e251-265, doi:10.1016/S2214-109X(16)00005-X (2016).

61 Stoltzfus, R. J. *et al.* Clinical pallor is useful to detect severe anemia in populations where anemia is prevalent and severe. *The Journal of nutrition* **129**, 1675-1681 (1999).

62 Strobach, S. R., Anderson, S. K., Doll, D. C. & Ringenberg, S. Q. The Value of the Physical Examination in the Diagnosis of Anemia: Correlation of the Physical Findings and the Hemoglobin Concentration. *Archives of Internal Medicine* **148**, 831-832, doi:10.1001/archinte.1988.00380040071013 (1988).

63 Kalter, H. D. *et al.* Evaluation of clinical signs to diagnose anaemia in Uganda and Bangladesh, in areas with and without malaria. *Bulletin of the World Health Organization* **75 Suppl 1**, 103-111 (1997).

64 Muhe, L., Oljira, B., Degefu, H., Jaffar, S. & Weber, M. W. Evaluation of clinical pallor in the identification and treatment of children with moderate and severe

anaemia. *Tropical Medicine & International Health* **5**, 805-810, doi:10.1046/j.1365-3156.2000.00637.x (2000).

65 Mogensen, C. B., Sørensen, J. E. & Bjorkman, A. Pallor as a sign of anaemia in small Tanzanian children at different health care levels. *Acta Tropica* **99**, 113-118, doi:10.1016/j.actatropica.2005.12.010 (2006).

66 Lin, J. Y. & Fisher, D. E. Melanocyte biology and skin pigmentation. *Nature* **445**, 843-850, doi:10.1038/nature05660 (2007).

67 Perrin, C., Michiels, J. F., Pisani, A. & Ortonne, J. P. Anatomic distribution of melanocytes in normal nail unit: an immunohistochemical investigation. *The American Journal of dermatopathology* **19**, 462-467 (1997).

68 Braun, R. *et al.* Diagnosis and management of nail pigmentations. *Journal of the American Academy of Dermatology* **56**, 835-847, doi:10.1016/j.jaad.2006.12.021 (2007).

69 Silva, B. M., Rodrigues, J. J., de la Torre Diez, I., Lopez-Coronado, M. & Saleem, K. Mobile-health: A review of current state in 2015. *J Biomed Inform* **56**, 265-272, doi:10.1016/j.jbi.2015.06.003 (2015).

70 Software, M. *Daily Carb - Carbohydrate, glucose, medication, blood pressure and excersize tracker*, <https://itunes.apple.com/us/app/daily-carbcarbohydrate-glucose/id536425111?mt=8> (2018).

71 ZocDoc. *ZocDoc - Doctor appointments online.*, <https://itunes.apple.com/us/app/zocdoc-doctor-appointments/id391062219?mt=8> (2018).

72 WebMD. *WebMD*, <https://itunes.apple.com/us/app/webmd-foripad/id373185673?mt=8> (2018).

73 Bonamici, S. (ed United States Congress) (2016).

74 Research2Guidance. mHealth App Economics 2017: Current status and future trends in mobile health (2017).

75 SkinVision. *SkinVision*, <https://www.skinvision.com> (2018).

76 Ayache, J. S. D. a. N. Medical image analysis: progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, 85-106, doi:10.1109/34.824822 (2000).

77 Cheng, J.-Z. *et al.* Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Scientific Reports* **6**, 24454, doi:10.1038/srep24454 (2016).

78 Greenspan, H. & van Ginneken, B. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on …* (2016).

79 Balarajan, Y., Ramakrishnan, U., Ozaltin, E., Shankar, A. H. & Subramanian, S. V. Anaemia in low-income and middle-income countries. *Lancet* **378**, 2123-2135, doi:10.1016/S0140-6736(10)62304-5 (2011).

80 Yager, P., Domingo, G. J. & Gerdes, J. Point-of-care diagnostics for global health. *Annu. Rev. Biomed. Eng.* (2008).

81 Kohli-Kumar, M. Screening for anemia in children: AAP recommendations--a critique. *Pediatrics* **108**, E56 (2001).

82 Needleman, H. Lead Poisoning. *Annual Review of Medicine* **55**, 209-222, doi:10.1146/annurev.med.55.091902.103653 (2004).

83 Zucker, J. R. *et al.* Clinical signs for the recognition of children with moderate or severe anaemia in western Kenya. *Bulletin of the World Health Organization* **75 Suppl 1**, 97-102 (1997).

84 Kwiatkowska, M., Jakutowicz, T., Ciszek, B. & Czubak, J. Can palmar creases serve as landmarks for the deeper neuro-vascular structures? *Surg Radiol Anat* **36**, 495-501, doi:10.1007/s00276-013-1211-4 (2014).

85 Bickley, L. S. *Bate's Guide to Physical Examination and History Taking*. 8 edn, (Lippincott Williams & Wilkins, 2003).

86 Hernández-Martín, A., Ros-Forteza, S. & de Unamuno, P. Longitudinal, transverse, and diffuse nail hyperpigmentation induced by hydroxyurea. *Journal of the American Academy of Dermatology* **41**, 333-334 (1999).

87 Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical …*, doi:10.1080/01621459.1979.10481038 (1979).

88 Hawkins, D. M. The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences* **44**, 1-12, doi:10.1021/ci0342472 (2004).

89 Niel, T. G., McVicar, T. R. & Datt, B. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. *Remote Sensing of Environment* **98**, 468-480, doi:10.1016/j.rse.2005.08.011 (2005).

90 Friston, K. J. *et al.* Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* **2**, 189-210, doi:10.1002/hbm.460020402 (1994).

91 Chumney, E. C. G. & Simpson, K. N. Methods and Designs for Outcomes Research. *American Society of Health Systems Pharmacists* (2006).

92 Skelton, V. A. *et al.* Evaluation of point-of-care haemoglobin measuring devices: a comparison of Radical-7™ pulse co-oximetry, HemoCue® and laboratory haemoglobin measurements in obstetric patients*. *Anaesthesia* **68**, 40-45, doi:10.1111/anae.12039 (2013).

93 Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. *Statistical methods in medical research*, doi:10.1177/096228029900800204 (1999).

94 Mortimer, R. J. & Varley, T. S. Quantification of colour stimuli through the calculation of CIE chromaticity coordinates and luminance data for application to in situ colorimetry studies of electrochromic materials. *Displays* **32**, 35-44, doi:10.1016/j.displa.2010.10.001 (2011).

95 Takiwaki, H. Measurement of skin color: practical application and theoretical considerations. *The journal of medical investigation : JMI* **44**, 121-126 (1998).

96 Srivastava, T., Negandhi, H., Neogi, S. B. & Sharma, J. Methods for hemoglobin estimation: a review of "what works.". *J Hematol* … (2014).

97 Prasad, A. M., Iverson, L. R. & Liaw, A. Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems* **9**, 181-199, doi:10.1007/s10021-005-0054-1 (2006).

98 Organization, W. Iron deficiency anaemia: assessment, prevention and control: a guide for programme managers. *Iron deficiency anaemia: assessment, prevention and control: a guide for programme managers* (2001).

99 Knutson, T. *et al.* Evaluation of a new nonnvasive device in determining hemoglobin levels in emergency department patients. *The western journal of emergency medicine* **14**, 283-286, doi:10.5811/westjem.2011.9.6733 (2013).

100 Radtke, H., Polat, G., Kalus, U., Salama, A. & Kiesewetter, H. Hemoglobin screening in prospective blood donors: Comparison of different blood samples and different quantitative methods. *Transfusion and Apheresis Science* **33**, 31-35, doi:10.1016/j.transci.2004.11.004 (2005).

101 Bland, J. M. & Altman, D. J. Regression analysis. *Lancet* **1**, 908-909, doi:10.1016/S0140-6736(86)91008-1 (1986).

102 for & for, H. H. S. Medicare, Medicaid, and CLIA programs; laboratory requirements relating to quality systems and certain personnel qualifications. Final rule. *Federal register* **68**, 3639-3714 (2003).

103 Specht, D. F. A general regression neural network. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council* **2**, 568-576, doi:10.1109/72.97934 (1991).

104 Nemati, S. & Clifford, G. Detecting Heart Failure Using Wearables: A Pilot Study. *International Society for Computerized Electrocardiology* (0400).

105 Lossius, K., Eriksen, M. & Walløe, L. Fluctuations in blood flow to acral skin in humans: connection with heart rate and blood pressure variability. *The Journal of Physiology* **460**, 641-655, doi:10.1113/jphysiol.1993.sp019491 (1993).

106 Rubinstein, E. H. & Sessler, D. I. Skin-surface temperature gradients correlate with fingertip blood flow in humans. *Anesthesiology* (1990).

107 Yoruk, A. & Yukselgungor, H. Chemotherapy induced transverse leukonychia in children. *Int J Dermatol* **42**, 468-469 (2003).

108 Canny, J. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-8**, 679-698, doi:10.1109/TPAMI.1986.4767851 (1986).

109 Watson, R. T. Africa's contributions to information systems. *The African Journal of Information …* (2013).

110 Guidelines for the clinical use of red cell transfusions. *British Journal of Haematology* **113**, 24-31, doi:10.1046/j.1365-2141.2001.02701.x (2001).

111 Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115-118, doi:10.1038/nature21056 (2017).

112 McGann, P. T. *et al.* An accurate and inexpensive color-based assay for detecting severe anemia in a limited-resource setting. *American Journal of Hematology* **90**, 1122-1127, doi:10.1002/ajh.24180 (2015).

113 McMurdy, J. W., Jay, G. D., Suner, S. & Crawford, G. Noninvasive optical, electrical, and acoustic methods of total hemoglobin determination. *Clinical chemistry* **54**, 264-272, doi:10.1373/clinchem.2007.093948 (2007).

114 Armstrong, G. R., Dewey, C. E. & Summerlee, A. J. S. Iron release from the Lucky Iron Fish™: safety considerations. *Iron release from the Lucky Iron Fish™: safety considerations*.